Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Journal of Discrete Algorithms 2 (2004) 161–206

JOURNAL OF DISCRETE ALGORITHMS

www.elsevier.com/locate/jda

A polynomial time upper bound for the number of contacts in the HP-model on the face-centered-cubic lattice (FCC)

Rolf Backofen

Institut für Informatik, FSU Jena, Ernst-Abbe-Platz 1-4, D-07743 Jena, Germany

Abstract

Lattice protein models are a major tool for investigating principles of protein folding. For this purpose, one needs an algorithm that is guaranteed to find the minimal energy conformation in some lattice model (at least for some sequences). So far, there are only algorithm that can find optimal conformations in the cubic lattice. In the more interesting case of the face-centered-cubic lattice (FCC), which is more protein-like, there are no results. One of the reasons is that for finding optimal conformations, one usually applies a branch-and-bound technique, and there are no reasonable bounds known for the FCC. We will give such a bound for Dill's HP-model on the FCC, which can be calculated by a dynamic programming approach.

Keywords: Protein folding; Lattice proteins; HP-model; Face-centered-cubic-lattice (FCC); Branch-and-bound

1. Introduction

Simplified protein models such as lattice models are used to investigate the protein folding problem, the major unsolved problem in computational biology. An important representative of lattice models is the HP-model, which has been introduced by [8]. In this model, the 20 letter alphabet of amino acids (called monomers) is reduced to a two letter alphabet, namely H and P. H represents *hydrophobic* monomers, whereas P represent *polar* or hydrophilic monomers. A *conformation* is a self-avoiding walk on the cubic lattice. The energy function for the HP-model simply states that the energy contribution of a contact between two monomers is -1 if both are H-monomers, and 0 otherwise. Two monomers

E-mail address: backofen@informatik.uni-jena.de (R. Backofen).

1570-8667/$ – see front matter © 2003 Published by Elsevier B.V.
doi:10.1016/S1570-8667(03)00076-5

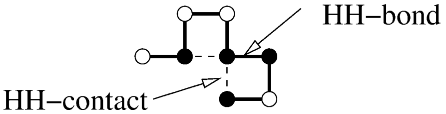


Fig. 1. Sample conformation.

form a *contact* in some specific conformation if they are not connected via a *bond*, and the euclidean distance of the positions is 1. One searches for a conformation which maximizes the number of contacts, which is a conformation whose hydrophobic core has minimal surface. Just recently, the structure prediction problem has been shown to be NP-hard even for the HP-model [4,6] on the cubic lattice. A sample conformation for the sequence PHPPHHPH in the two-dimensional lattice with energy -2 is given in Fig. 1. The white beads represent P, the black ones H monomers. The two contacts are indicated via dashed lines.

For investigating general properties of protein-folding, one needs an algorithm which is guaranteed to find a conformation with maximal number of contacts (at least for some sequences, since the problem is NP-hard in general). Although there are approximation algorithms for the HP-model in the cubic lattice [7] and FCC [1], the need of an optimal conformation in this case implies that one cannot use approximate or heuristic algorithms for this purpose. To our knowledge, there are two algorithms known in the literature that find conformations with maximal number of contacts (optimal conformations) for the HP-model, namely [2,11]. Both use some variant of Branch-and-Bound.

The HP-model is original defined for the cubic lattice, but it is easy to define it for any other lattice. Of special interest is the face-centered-cubic lattice (FCC), which models protein backbone conformations more appropriately than the cubic lattice. When considering the problem of finding an optimal conformation, the problem occurs that no good bound on the number of contacts for the face-centered cubic lattice is known, in contrast to the HP-model. Both known algorithm for finding the optimal conformation search through the space of conformations using the following strategy:

- fix one coordinate (say x) of all H-monomers first,
- calculate an upper bound on the number of contacts, given fixed values for the H-monomers.

An upper bound can easily be given in the case of the HP-model, if only the number of H-monomers are known in every plane defined by an equation $x = c$ (called x-layer in the following). For this purpose, one counts the number of HH-contacts and HH-bonds (since the number of HH-bonds is constant, and we do not care in which layer the HH-bonds actually are). Let us call this generalized contacts in the following. Then one distinguishes between generalized contacts within an x-layer, and generalized contacts between x-layers. Suppose that the positions occupied by H-monomers are given by black dots in Fig. 2. Then we have 5 H-monomers in layer $x = 1$, and 4 H-monomers in $x = 2$. Furthermore, we have 4 generalized contacts between the layer $x = 1$ and $x = 2$ (straight lines), 5 contacts within $x = 1$ and 4 contacts within $x = 2$ (dashed lines). This coincide with the upper bound given 5 H-monomer in $x = 1$, and 4 H-monomers in $x = 2$, which is calculated as follows. For the number of interlayer contacts, we know that every interlayer contact

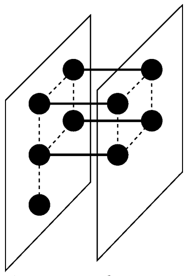


Fig. 2. H-positions.

consumes 1 H-monomer in every layer. Hence, the maximal number of interlayer contacts is the minimum of the number of H-monomer in each layer, in this case $\min(5,4)=4$. The upper bound for the layer contacts is a bit more complicated, since it uses the concept of a frame. Consider some layer with n H-monomers. Let $a=\lceil\sqrt{n}\,\rceil$ and $b=\lceil\frac{n}{a}\rceil$. (a,b) is the minimal rectangle (frame) around n H-monomers. Then the maximal number of contacts within this layer is upper bound by $2n-a-b$. In our example, we get for the first layer $n=5$, $a=3$ and $b=2$, and the maximal number of layer contacts is then $10-3-2=5$, as it is the case in our example. For $n=4$, we get $a=2$, $b=2$ and the maximal number is then $8-2-2=4$, as in our case. For details, see [2,11].

For the face-centered-cubic lattice (FCC) is no similar bound known, and there is no trivial transfer from the cubic lattice. The bound for FCC lattice is harder, since the inter-layer contacts are much more complex to determine. The reason is that the FCC has 12 neighbors (position with minimal distance), whereas the cubic lattice has only 6. Thus, in any representation of FCC, we have more than one neighbor in the next x-layer for any point $\vec{p}$, which makes the problem complicated. Such an upper bound will be given in this paper.

2. Preliminaries

Given vectors $\vec{v}_1,\ldots,\vec{v}_n$, the lattice generated by $\vec{v}_1,\ldots,\vec{v}_n$ is the minimal set of points L such that $\forall\vec{u},\vec{v}\in L$, both $\vec{u}+\vec{v}\in L$ and $\vec{u}-\vec{v}\in L$. An *x-layer* in a lattice L is a plane orthogonal to the x-axis (i.e., is defined by the equation $x=c$) such that the intersection of the points in the plane and the points of L is non-empty. The *face-centered cubic lattice* (short FCC, see [5]) is defined as the lattice

$$D_3=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\;\middle|\;\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{Z}^3 \text{ and } x+y+z \text{ is even}\right\}.$$

For simplicity, we use a representation of D_3 that is rotated by $\phi=45^\circ$ along the x-axis. Since we want to have distance 1 between successive x-layers, and unit distance between neighbors in one x-layer, we additionally scale the y- and z-axis, but leave the x-axis as it is. A partial view of the lattice and its connections, as well as the rotated lattice is given in Fig. 3. Thus, we can define the lattice D_3' to be the lattice that consists of the following

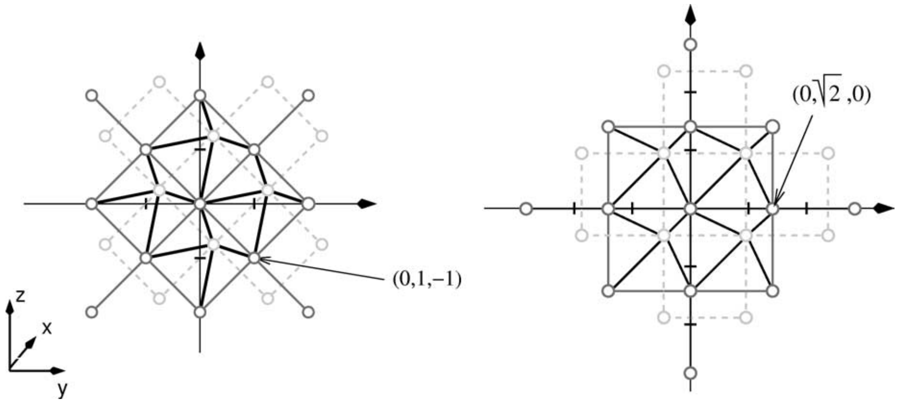


Fig. 3. In the first figure, we have shown two x-layers (where the x-axis is shown as the third dimension). The dark grey circles are the lattice points in the first x-layer (where the dark grey, solid lines are the nearest neighbor connections). The light grey circles are the points in the second x-layers (where the light grey, dashed lines are the nearest neighbor connections in the second layer). The black lines indicate the nearest neighbor connections between the first and the second x-layer. The second figure shows FCC after rotation by 45°.

sets of points in real coordinates:

$$D_3' = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{Z}^3 \text{ and } x \text{ is even} \right\}$$
$$\uplus \left\{ \begin{pmatrix} x \\ y+0.5 \\ z+0.5 \end{pmatrix} \;\middle|\; \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{Z}^3 \text{ and } x \text{ odd} \right\}.$$

The first is the set of points in even x-layers, the second the set of point in odd x-layers. A generator matrix for D_3' is given in [3].

The set $N_{D_3'}$ of minimal vectors connecting neighbors in D_3' is given by

$$N_{D_3'} = \left\{ \begin{pmatrix} 0 \\ \pm 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \pm 1 \end{pmatrix} \right\} \uplus \left\{ \begin{pmatrix} \pm 1 \\ \pm 0.5 \\ \pm 0.5 \end{pmatrix} \right\}.$$

The vectors in the second set are the vectors connecting neighbors in two different successive x-layers. Two points $\vec{p}$ and $\vec{p}'$ in D_3' are *neighbors* if $\vec{p} - \vec{p}' \in N_{D_3'}$.

2.1. Colorings

We are interested in the positions occupied by H-monomers in some conformation of the HP-model. For this purpose, we introduce the concept of *colorings* (where the *colored* points are the points occupied by H-monomers).

Definition 2.1 (*Coloring*). A *coloring* is a function $f : D_3' \to \{0, 1\}$. We denote with $\text{points}(f)$ the set of all points colored by f, i.e.,

$$\text{points}(f) = \left\{ \vec{p} \mid f(\vec{p}) = 1 \right\}.$$

With $\mathrm{num}(f)$ we denote $|\mathrm{points}(f)|$. Let f_1 and f_2 be colorings. With $f_1 \cup f_2$ we denote the coloring f with

$$\mathrm{points}(f) = \mathrm{points}(f_1) \cup \mathrm{points}(f_2).$$

Two colorings f_1, f_2 are *disjoint* if their set of points are disjoint. $f_1 \uplus f_2$ denotes the disjoint union of colorings. Given a coloring f, we define the *number of contacts* $\mathrm{con}(f)$ *of* f by $\mathrm{con}(f) = \frac{1}{2}|\{(\vec{p}, \vec{p}') \mid f(\vec{p}) = 1 = f(\vec{p}') \wedge (\vec{p} - \vec{p}') \in N_{D_3'}\}$.

In the following, we will split a complete coloring f into its composition of colorings of the single x-layers that contain points colored by f. The aim is to give separate bounds for layer and interlayer contacts. For this purpose, we introduce colorings, where the colored points are contained in one x-layer.

Definition 2.2 (*Plane coloring*). A coloring f is called a *coloring of the plane* $x = c$ if $f(x, y, z) = 1$ implies $x = c$. We say that f is a *plane coloring* if there is a c such that f is a coloring of plane $x = c$. We define $\mathrm{Surf}_{pl}(f)$ to be the surface of f in the plane $x = c$, i.e.,

$$\mathrm{Surf}_{pl}(f) = \Bigg\{(\vec{p}, \vec{p}') \mid (\vec{p} - \vec{p}') \in N_{D_3'} \wedge f(\vec{p}) = 1 \wedge \\ f(\vec{p}') = 0 \wedge \exists y, z\colon\ \vec{p}' = \begin{pmatrix} c \\ y \\ z \end{pmatrix}\Bigg\}.$$

With $\min_y(f)$ we denote the integer

$$\min\big\{y \mid \exists z\colon\ f(c, y, z) = 1\big\}.$$

$\max_y(f)$, $\min_z(f)$ and $\max_z(f)$ are defined analogously.

$\min_y(f)$, $\max_y(f)$, $\min_z(f)$ and $\max_z(f)$ defines the minimal rectangle that contains all points colored by the plane coloring f.

3. Description of the upper bound

Our purpose is to give an upper bound on the number of contacts, given that n_c H-monomers are in the x-layer defined by $x = c$. Thus, we need to find a function $b(n_1, \ldots, n_k)$ with

$$b(n_1, \ldots, n_k) \geqslant \max\left\{ \mathrm{con}(f) \,\middle|\, \begin{array}{l} f = f_1 \uplus \cdots \uplus f_k, \forall c \in \{1, \ldots, k\}\colon\ f_c \text{ is a} \\ \text{coloring of plane } x = c \text{ and } \mathrm{num}(f_c) = n_c \end{array} \right\}.$$

To develop $b(n_1, \ldots, n_k)$, we distinguish between contacts $(\vec{p}, \vec{p}')$ where both $\vec{p}$ and $\vec{p}'$ are in one x-layer, and contacts $(\vec{p}, \vec{p}')$ where $\vec{p}$ is in an layer $x = c$, and $\vec{p}'$ is in the layer $x = c + 1$. The contacts within the same x-layer are easy to bound by bounding the surface $\mathrm{Surf}_{pl}(f_c)$. Since every point in layer $x = c$ has four neighbors, which are either occupied

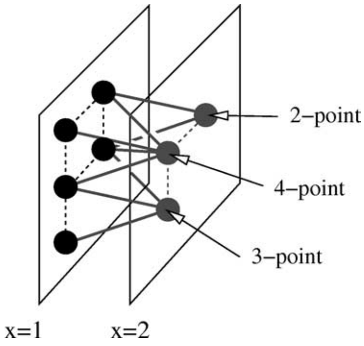


Fig. 4. H-positions in FCC.

by an colored point, or an uncolored point, we get $4 \cdot \mathrm{num}(f) = \mathrm{Surf}_{pl}(f_c) + 2 \cdot LC$, where LC is the number of layer contacts. The hard part is to bound the number of contacts between two successive layers.

For defining the bound on the number of contacts between two successive layers, we introduce the notion of an i-point, where $i = 1, 2, 3, 4$. Given any point in $x = c + 1$, then this point can have at most 4 neighbors in the plane $x = c$. Let f be a coloring of the plane $x = c$. Then a point $\vec{p}$ in plane $x = c + 1$ is an *i-point for f* if it has i neighbors in plane $x = c$ that are colored by f (where $i \leqslant 4$). Of course, if one colors an i-point in plane $x = c + 1$, then this point generates i contacts between layer $x = c$ and $x = c + 1$. In the following, we will restrict ourself to the case where $c = 1$ for simplicity. Of course, the calculation is independent of the choice of c.

Consider as an example the two colorings f_1 of plane $x = 1$ and f_2 of plane $x = 2$ as shown in Fig. 4. f_1 consists of 5 colored points, and f_2 of 3 colored points. Since f_2 colors one 4-point, one 3-point and one 2-point of f_1, there are 9 contacts between these two layers. It is easy to see that we generated the most contacts between layers $x = 1$ and $x = 2$ by first coloring the 4-points, then the 3 points and so on until we reach the number of points to be colored in layer $x = 2$.[1]

For this reason, we are interested to calculate the maximal number of i-points (for $i = 1, 2, 3, 4$), given only the number of colored points n in layer $x = 1$. But this would overestimate the number of possible contacts, since we would maximize the number of 4-, 3-, 2- and 1-point independently from each other. We have found a dependency between these numbers, which requires to fix the side length (a, b) of the minimal rectangle around all colored points in layer $x = 1$ (called the *frame*). In our example, the frame is $(3, 2)$. Of course, one has to search through all "reasonable frames" to find the maximal number of contacts between the two layers. This will be treated in a later section.

[1] Note that this might not necessarily be the coloring with the maximal number of contacts, since we might loose contacts within the layer $x = 2$; although this could be included in the calculation of the upper bound, we have excluded this effect for simplicity.

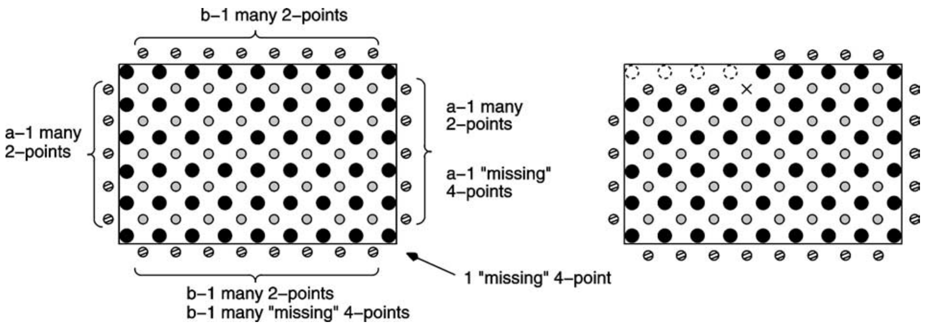


Fig. 5. Colorings and corresponding 4-, 3- and 2-points (1-points are not shown). 4-points are indicated by ◯, 2-points by ⊘, and the single 3-point by a $\times$.

Denote with $max_i(a,b,n)$ the maximal number of i-points in layer $x=2$ for any coloring of layer $x=1$ with n-colored points and frame (a,b). Then we have found that

$$max_4(a,b,n)=n+1-a-b, \qquad max_2(a,b,n)=2a+2b-2\ell-4,$$
$$max_3(a,b,n)=\ell, \qquad max_1(a,b,n)=\ell+4.$$

The remaining part is to find $\ell = max_3(a,b,n)$, which is a bit more complicated.

Before we will do so, let us explain $max_4(a,b,n)$ and $max_2(a,b,n)$ first. Consider the left coloring in Fig. 5, which is a coloring that completely fills its frame (with $a=6$ and $b=9$). This coloring contains $n=54$ points. If one shifts this n colored points by $(1,-0.5,0.5)$, than one gets all 4-points except the $a-1$ "missing" 4-points in the bottom row, the $b-1$ "missing" 4-points in the last column, and the one "missing" 4-point in the right bottom corner. This makes

$$n-(a-1)-(b-1)-1=n+1-a-b$$

as given by $max_4(a,b,n)$. For the 2-points, we have $2a+2b-4$ many 2-points, where the -4 stems from the "missing" 2-points at the 4 corners (which are in fact 1-points).

Now the interesting part is that basically, this relation does not change if we remove some colored points. Consider the right coloring in Fig. 5, which has four colored points deleted. By removing four colored points, we remove four 4-points. Hence, we have again that the number of 4-points is $n+1-a-b$. For the 2-points, four 2-points have been deleted in the top row, and one additional 2-point has been deleted in the first column. But the four deleted 4-points now have become 2-points except one, which has become a 3-point. One could say that the 3-point has been generated by merging two moved 2-points (one from the top row, and one from the first column). Hence, we have that the number of 2-points is

$$2a+2b-4-2\ell,$$

where ℓ is the number of 3-points.

Now let's turn to the 3- and 1-points. For finding $\ell = max_3(a,b,n)$, we define

$$k=\text{edge}(a,b,n)=\max\left\{k\in\mathbb{N} \mid ab-4\frac{k(k+1)}{2}\geqslant n\right\},$$

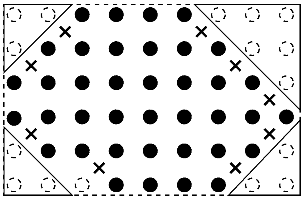


Fig. 6. Uncolored diagonals and 3-points, which are indicated by $\times$.

and

$$r = \mathrm{ext}(a, b, n) = \left\lfloor \frac{ab - 4\frac{k(k+1)}{2} - n}{k+1} \right\rfloor.$$

The geometric interpretation of $k = \mathrm{edge}(a, b, n)$ and $r = \mathrm{ext}(a, b, n)$ is the following. k is the maximal number of diagonals that can be left uncolored in all corners of the frame (when distributing the uncolored diagonals equally on all corners). $r \leqslant 3$ is the number of times that we can add one additional uncolored diagonal.

To give an example, consider the coloring in Fig. 6 with $n = 38$, $a = 6$ and $b = 9$. Then $k = \mathrm{edge}(a, b, n)$ is 2. That means, that in each corner we can have at least 2 diagonal lines that are uncolored. $r = \mathrm{edge}(a, b, n)$ is 1, which means that in one corner, we can add a third uncolored diagonal.

Now the interesting part is, that the number of uncolored diagonal determines the number of 3-points. Consider the left upper corner. There are two uncolored lines, and two 3-points are generated in this corner. The same relation holds for all other corners as well. We will show that we can define the bound on the number of 3-points by

$$max_3(a, b, n) = \begin{cases} 4k + r & \text{if } 4k + r < 2(a - 1), \\ 2(a - 1) & \text{else} \end{cases}$$

(assuming without loss of generality that $a \leqslant b$).

For the number of 1-points, it is easy to see that every corner produces one 1-point. For every 3-point, one additional 1-point is generated, which gives $\ell + 4$, where ℓ is the number of 3-points.

3.1. Plan of the paper

In Section 4, we will determine the number of points having n possible contacts, given some parameter of the coloring f of plane $x = c$. The parameters are the surface $\mathrm{Surf}_{pl}(f)$, and the number of points with 3 possible contacts.

In Section 5, we will then show how we can determine the number of points having 3 possible contacts, given $\mathrm{Surf}_{pl}(f)$. $\mathrm{Surf}_{pl}(f)$ is determined by the minimal rectangle (called frame) around all points colored by f. Thus, we get an upper bound for both the contacts in the plane $x = c$, and the contacts between $x = c$ and $x = c + 1$ by enumerating all possible frames for f. Of course, we cannot enumerate *all* frames. Thus, we introduce in Section 6 a concept of “sufficiently filled frames”, i.e., frames that are not too big for the number of points to be colored within the frame. These frames will be called normal.

Then, we prove that it is sufficient to enumerate only the normal frames to get an upper bound. In fact, this is the most tedious part of the construction. In Section 7, we combine the results in a dynamic programming approach, which allows to calculate the upper bound for the number of contacts in polynomial time. We will compare our bound with the trivial $6n$ bound used so far in the literature.

4. Number of points with 1, 2, 3, 4-contacts

In the following, we want to handle caveat-free, connected colorings, which we will define first.

Definition 4.1 (*Path, connected coloring*). Let f be a coloring of the plane $x = c$, and let $\vec{p}$ and $\vec{p'}$ be two points such that $f(\vec{p}) = 1 = f(\vec{p'})$. A *path between* $\vec{p}$ *and* $\vec{p'}$ *in* f is a list of points

$$\vec{p} = \vec{p}_1 \dots \vec{p}_n = \vec{p'}$$

such that

$$\forall 1 \leqslant i < n: \quad (\vec{p_{i+1}} - \vec{p}_i) \in \left\{ \pm \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \pm \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \pm \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \pm \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

A coloring f is *connected* if for any two points $\vec{p}$ and $\vec{p'}$ with $f(\vec{p}) = 1 = f(\vec{p'})$, there is a path between $\vec{p}$ and $\vec{p'}$ in f.

Definition 4.2 (*Caveats*). Let f be a coloring of plane $x = c$. A *horizontal caveat in* f is a k-tuple of points $(\vec{p}_1, \dots, \vec{p}_k)$ such that

$$\begin{aligned}
&\forall 1 \leqslant j < k: \quad \left(\vec{p_{j+1}} = \vec{p}_j + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right), \\
&f(\vec{p}_1) = 1 = f(\vec{p}_k), \\
&\forall 1 < j < k: \quad f(\vec{p}_j) = 0.
\end{aligned}$$

A *vertical caveat in* f is defined analogously satisfying

$$\forall 1 \leqslant j < k: \quad \left(\vec{p_{j+1}} = \vec{p}_j + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

instead. We say that f contains a *caveat* if there is at least one horizontal or vertical caveat in f. f is called *caveat-free* if it does not contain a caveat.

For calculating the number of contacts, we distinguish for a plane coloring f the points in the next and previous plane according to the number of contacts that can be achieved by coloring the specific point.

Definition 4.3. Let f be a coloring of plane $x = c$. We say that a point $\vec{p}$ is a 4-point for f if $\vec{p}$ is in plane $x = c + 1$ or $x = c - 1$ and $\vec{p}$ has 4 neighbors $\vec{p}_1, \ldots, \vec{p}_4$ in plane $x = c$ with $f(\vec{p}_1) = \cdots = f(\vec{p}_4) = 1$. Analogously, we define 3-points, 2-points and 1-points. Furthermore, we define $\#4_{c-1}(f) = |\{\vec{p} \mid \vec{p}$ is a 4-point for f in $x = c - 1\}|$. Analogously, we define $\#4_{c+1}(f)$ and $\#i_{c\pm1}(f)$ for $i = 1, 2, 3$.

Trivially, we get for any coloring f of plane $x = c$ that $\forall i \in [1..4]$: $\#i_{c-1}(f) = \#i_{c+1}(f)$. Hence, we define for a coloring f of plane $x = c$ that $\#i(f) = \#i_{c-1}(f)$ $(= \#i_{c+1}(f))$ for every $i \in [1..4]$. For calculating the number of i-points for a coloring f of plane $x = c$, we need the additional notion of an x-steps for f. An x-step f consists of 3 points in $x = c$ that are sufficient to characterize one 3-point. Furthermore, we need to now whether the lines of the coloring overlap or not.

Definition 4.4 (*X-step*). Let f be a coloring of plane $x = c$. An *x-step for f* is a triple $(\vec{p}_1, \vec{p}_2, \vec{p}_3)$ such that

$$
\begin{aligned}
&f(\vec{p}_1) = 0,\\
&f(\vec{p}_2) = 1 = f(\vec{p}_3),\\
&\vec{p}_1 - \vec{p}_2 = \pm \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\\
&\vec{p}_1 - \vec{p}_3 = \pm \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.
\end{aligned}
$$

With $\mathrm{xsteps}(f)$ we denote the number of x-steps of f.

Definition 4.5 (*Overlaps*). Let f be a coloring of plane $x = c$. We define

$$
\begin{aligned}
&\text{r\_overlap}^+(f, z) = \left| \left\{ y \,\middle|\, \begin{array}{l} f(c, y, z) = 1 \wedge f(c, y, z+1) = 1 \\ \wedge\, \exists y(f(c, y, z) = 1) \wedge \exists y(f(c, y, z+1) = 1) \end{array} \right\} \right|,\\
&\#\text{r\_not\_overlaps}(f) = \left| \left\{ z \mid \min_z(f) \leqslant z < \max_z(f) \wedge \text{r\_overlap}^+(f, z) = 0 \right\} \right|.
\end{aligned}
$$

Lemma 4.6. *Let f be a connected, horizontal caveat-free coloring of the plane $x = c$. Then the following equations are valid*:

$$
\begin{aligned}
\#4(f) &= \mathrm{num}(f) + 1 - \frac{1}{2}\mathrm{Surf}_{pl}(f) + \#\text{r\_not\_overlaps}(f), && (1)\\
\#3(f) &= \mathrm{xsteps}(f) - 2\#\text{r\_not\_overlaps}(f), && (2)\\
\#2(f) &= 2\mathrm{num}(f) - 2\#4(f) - 2\#3(f) - 2 - \#\text{r\_not\_overlaps}(f), && (3)\\
\#1(f) &= \#3(f) + 4 + 2\#\text{r\_not\_overlaps}(f) && (4)\\
&= \mathrm{xsteps}(f) + 4.
\end{aligned}
$$

Proof. Claims (1), (2) and (4) are proven by induction on the height of f.

Base case. For the base case that f has height 1, we know that we have $\#4(f) = 0$, $\#3(f) = 0$, $\#1(f) = 4$ and that $\mathrm{Surf}_{pl}(f) = 2n + 2$. Thus, claims (1), (2) and (4) hold.

Induction step. Let f be a plane coloring of height $h + 1$. Let the coloring f' be f with the row $z = \max_z(f)$ deleted.

Claim (1). Let n_r be the number of points introduced in the last line $z = \max_z(f)$ in f. Let $r = \mathrm{r\_overlap}^+(f, \max_z(f) - 1)$. We have two cases:

1. $r = 0$. This implies $\mathrm{Surf}_{pl}(f) = \mathrm{Surf}_{pl}(f') + 2n_r + 2$. Furthermore, $\#4(f) = \#4(f')$ and

$$\#\mathrm{r\_not\_overlaps}(f) = \#\mathrm{r\_not\_overlaps}(f') + 1.$$

Thus we get by induction hypotheses

$$\begin{aligned}\#4(f) &= \#4(f')\\ &= \mathrm{num}(f') + 1 - \frac{1}{2}\mathrm{Surf}_{pl}(f') + \#\mathrm{r\_not\_overlaps}(f')\\ &= (\mathrm{num}(f) - n_r) + 1 - \frac{1}{2}\big(\mathrm{Surf}_{pl}(f) - 2n_r - 2\big)\\ &\quad + \big(\#\mathrm{r\_not\_overlaps}(f) - 1\big)\\ &= \mathrm{num}(f) + 1 - \frac{1}{2}\mathrm{Surf}_{pl}(f) + \#\mathrm{r\_not\_overlaps}(f).\end{aligned}$$

2. $r > 0$. This implies $\mathrm{Surf}_{pl}(f) = \mathrm{Surf}_{pl}(f') + 2(n_r - r) + 2$. Furthermore, $\#4(f) = \#4(f') + (r - 1)$ and $\#\mathrm{r\_not\_overlaps}(f) = \#\mathrm{r\_not\_overlaps}(f')$. Thus we get by induction hypotheses:

$$\begin{aligned}\#4(f) &= \#4(f') + (r - 1)\\ &= \mathrm{num}(f') + 1 - \frac{1}{2}\mathrm{Surf}_{pl}(f')\\ &\quad + \#\mathrm{r\_not\_overlaps}(f') + (r - 1)\\ &= (\mathrm{num}(f) - n_r) + 1 - \frac{1}{2}\big(\mathrm{Surf}_{pl}(f) - 2(n_r - r) - 2\big)\\ &\quad + \#\mathrm{r\_not\_overlaps}(f) + (r - 1)\\ &= \mathrm{num}(f) - n_r + 1 - \frac{1}{2}\mathrm{Surf}_{pl}(f) + n_r - r + 1\\ &\quad + \#\mathrm{r\_not\_overlaps}(f) + r - 1\\ &= \mathrm{num}(f) - \frac{1}{2}\mathrm{Surf}_{pl}(f) + 1 + \#\mathrm{r\_not\_overlaps}(f).\end{aligned}$$

Claims (2) *and* (4). We have listed in Fig. 7 all cases of how the last two lines of f can overlap (or not). In any case where we have an overlap, then the introduction of an x-step between the last two lines yields in f an additional 3-point and an additional 1-point.

If there is no overlap between the last two lines, then there are two x-steps (since f is connected). But these introduce no additional 3-points, but two additional 1-points.

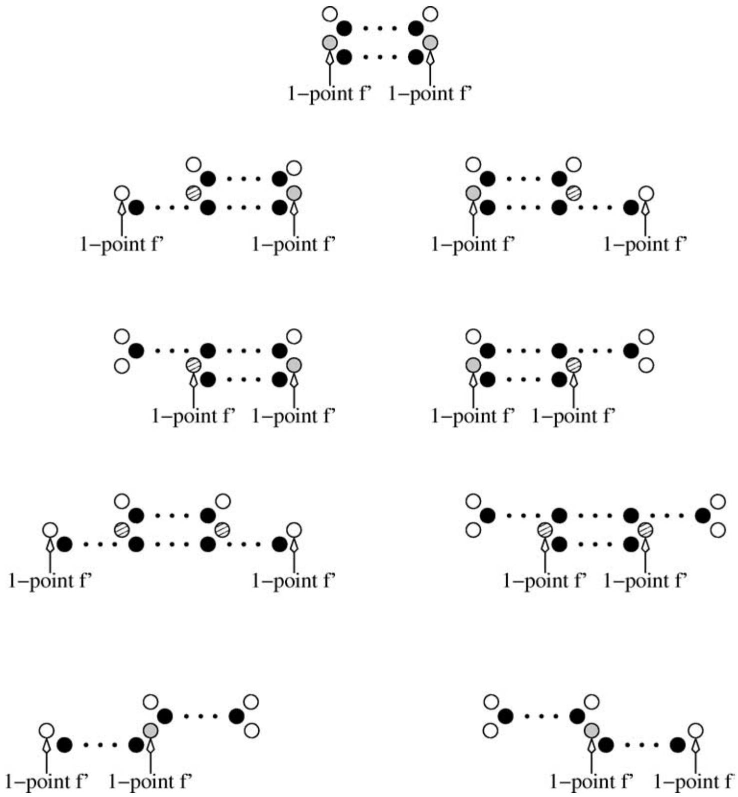


Fig. 7. Cases for claims (2) and (4). 1-points for f are indicated by , 2-points for f by , and 3-points for f by . We have indicated all 1-points for f', and have shown only those 2-points for f which have been 1-points for f'. Note that some of them are also 1-points for f, other change into a 2-point or 3-point for f.

For claim (3), we first note that the sum of contacts of all 4-, 3-, 2- and 1-points must yield $4n$, since this is the number of contacts that can be achieved if all those points are filled in the next plane. Hence,

$$\begin{aligned} 2\#2(f) &= 4n - 4\#4(f) - 3\#3(f) - 1\#1(f) \\ &= 4n - 4\#4(f) - 3\#3(f) - \big(\#3(f) + 4 + 2\#\text{r\_not\_overlaps}(f)\big) \\ &= 4n - 4\#4(f) - 4\#3(f) - 4 - 2\#\text{r\_not\_overlaps}(f). \end{aligned}$$

This gives

$$\#2(f) = 2n - 2\#4(f) - 2\#3(f) - 2 - \#\text{r\_not\_overlaps}(f). \qquad \square$$

We will show later in Lemma 6.12 that it is sufficient to consider only the case of plane colorings, where successive colored lines overlap. In principle, this lemma can be used to show that our bound is even valid for all caveat-free colorings (thus skipping the additional condition that the coloring must be connected), albeit this is not explicitly proven in this paper.

Corollary 4.7. *Let f be a coloring of the plane $x = c$ with the property that*

$$\#\mathrm{r\_not\_overlaps}(f) = 0.$$

Then

$$\#4(f) = n + 1 - \frac{1}{2}\mathrm{Surf}_{pl}(f), \tag{5}$$

$$\#3(f) = \mathrm{xsteps}(f), \tag{6}$$

$$\#2(f) = 2n - 2\#4(f) - 2\#3(f) - 2, \tag{7}$$

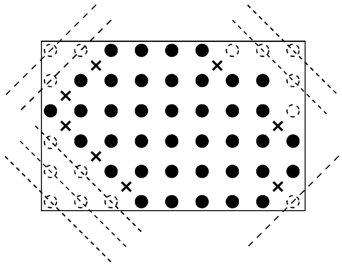
$$\begin{aligned} \#1(f) &= \#3(f) + 4 \\ &= \mathrm{xsteps}(f) + 4. \end{aligned} \tag{8}$$

With this corollary, we need only to bound $\mathrm{Surf}_{pl}(f)$ and $\mathrm{xsteps}(f)$ (which is a bound on the number of 3-points) to calculate bounds on the number of 4-, 3-, 2- and 1-points.

5. Bound on the number of 3-points

Given a plane coloring f, then we denote with $\mathrm{frame}(f)$ the pair (a, b), where $a = \max_z(f) - \min_z(f) + 1$ and $b = \max_y(f) - \min_y(f) + 1$. a is called the *height of* f, and b is called the *width of* f. The frame gives us the possibility to calculate a lower bound on the surface of a plane coloring, which is then an upper bound on the layer contacts. We need more information about a coloring than the frame to generate a bound for $\mathrm{xsteps}(f)$, which will be captured by the notion of a detailed frame. The formal definition will be given later. In principle, the detailed frame just counts for every corner, how many diagonals we can draw (starting from the corner) without touching a point that is colored by f. E.g., consider the following plane coloring f_{ex} given by the black dots:

f_{ex}:



Note that there are 8 positions in the next layer that are 3-points for this coloring. We have indicated these points with a $\times$. We can draw 3 diagonals from the left-lower corner, 2 from the left upper, 1 from the right lower, and 2 from the right upper corner. Note that the number of 3-points near every corner is exactly the same. We will prove this relationship later.

The detailed frame of a coloring f is the tuple $(a, b, i_{lb}, i_{lu}, i_{rb}, i_{ru})$, where (a, b) is the frame of f, and i_{lb} is the number of diagonals that can be drawn from the left-bottom corner. i_{lu}, i_{rb}, i_{ru} are defined analogously. For f_{ex}, the detailed frame is $(6, 9, 3, 2, 1, 2)$.

The interesting part is that the number of diagonals to be drawn gives an upper bound for the number of points to be colored (Proposition 5.4) *and* for the number of x-steps (Lemma 5.6).

Now we start with the formal definition of a detailed frame.

Definition 5.1 (*Corner, inbound vector*). Let f be a coloring of the plane $x = c$. The set of *corners* $C(f)$ of f is defined by

$$C(f) = \left\{ \begin{pmatrix} c \\ \min_y(f) \\ \min_z(f) \end{pmatrix}, \begin{pmatrix} c \\ \min_y(f) \\ \max_z(f) \end{pmatrix}, \begin{pmatrix} c \\ \max_y(f) \\ \min_z(f) \end{pmatrix}, \begin{pmatrix} c \\ \max_y(f) \\ \max_z(f) \end{pmatrix} \right\}.$$

We will call these corners c^f_{lb}, c^f_{lu}, c^f_{rb}, and c^f_{ru}, respectively. We omit f if it is clear from the context. We define for every corner $c \in C(f)$ the *inbound vector* $\mathrm{in}^f(c)$ of c in f by

$$\mathrm{in}^f(c_{lb}) = \begin{pmatrix} 0 \\ 0.5 \\ 0.5 \end{pmatrix}, \qquad \mathrm{in}^f(c_{lu}) = \begin{pmatrix} 0 \\ 0.5 \\ -0.5 \end{pmatrix},$$
$$\mathrm{in}^f(c_{rb}) = \begin{pmatrix} 0 \\ -0.5 \\ 0.5 \end{pmatrix}, \qquad \mathrm{in}^f(c_{ru}) = \begin{pmatrix} 0 \\ -0.5 \\ -0.5 \end{pmatrix}.$$

In the following, we consider lines (i.e., one-dimensional, affine subspaces $U + \vec{u}$ of $\mathbb{R}^3$, where $U = \mathrm{Lin}(\vec{v})$ is the linear, one-dimensional subspace generated by the vector $\vec{v}$). We are mainly considering lines which are either parallel to either the y-axis, or the z-axis, or which are diagonal in an x-layer. The diagonal ones are defined as the affine subspaces

$$\mathrm{Lin}\begin{pmatrix} 0 \\ 0.5 \\ 0.5 \end{pmatrix} + \vec{u} \quad \text{or} \quad \mathrm{Lin}\begin{pmatrix} 0 \\ 0.5 \\ -0.5 \end{pmatrix} + \vec{u}.$$

Considering the diagonal lines, then there is for every corner exactly one diagonal line which cuts the frame around a plane coloring f in exactly one point (namely the corner itself). This leads to the definition of a tangent. We say that a line $L = \mathrm{Lin}(\vec{u}) + \vec{v}$ *intersects* with a coloring f if there is a point $\vec{p} \in L$ such that $f(\vec{p}) = 1$.

Definition 5.2 (*Tangent*). Let f be a plane coloring with frame (a, b). We define the *tangent vector* $\mathrm{tavec}^f(c)$ of the corner $c \in C(f)$ of f by

$$\mathrm{tavec}^f(c_{lb}) = \begin{pmatrix} 0 \\ 0.5 \\ -0.5 \end{pmatrix}, \qquad \mathrm{tavec}^f(c_{lu}) = \begin{pmatrix} 0 \\ 0.5 \\ 0.5 \end{pmatrix},$$
$$\mathrm{tavec}^f(c_{rb}) = \begin{pmatrix} 0 \\ 0.5 \\ 0.5 \end{pmatrix}, \qquad \mathrm{tavec}^f(c_{ru}) = \begin{pmatrix} 0 \\ 0.5 \\ -0.5 \end{pmatrix}.$$

Then the *tangent* $\mathrm{ta}^f(c)$ in the corner $c \in C(f)$ of f is defined by

$$\mathrm{ta}^f(c_{lb}) = \mathrm{Lin}\big(\mathrm{tavec}^f(c_{lb})\big) + c_{lb}, \qquad \mathrm{ta}^f(c_{lu}) = \mathrm{Lin}\big(\mathrm{tavec}^f(c_{lu})\big) + c_{lu},$$
$$\mathrm{ta}^f(c_{rb}) = \mathrm{Lin}\big(\mathrm{tavec}^f(c_{rb})\big) + c_{rb}, \qquad \mathrm{ta}^f(c_{ru}) = \mathrm{Lin}\big(\mathrm{tavec}^f(c_{ru})\big) + c_{ru}.$$

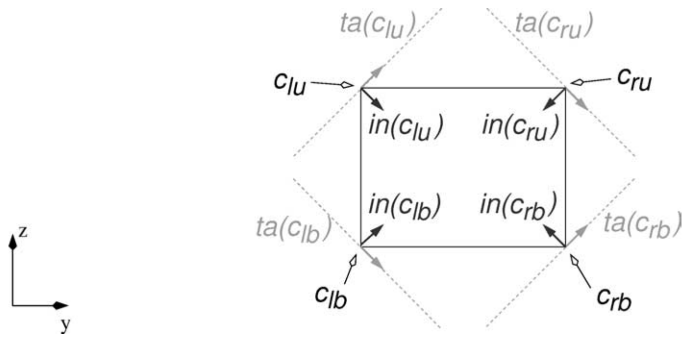


Fig. 8. Corner, inbound vectors and tangents. The inbound vectors are shown in dark grey, and the tangents in light grey.

Again, we omit the f if it is clear from the context. The above definitions are summarized in Fig. 8.

Definition 5.3 (*Detailed frame, characteristics*). The *detailed frame* of a plane coloring f with frame (a, b) is defined as the tuple $(a, b, i^f_{lb}, i^f_{lu}, i^f_{rb}, i^f_{ru})$, where i^f_j for $j \in \{lb, lu, rb, ru\}$ is the minimal integer such that $\mathrm{ta}(c_j) + i^f_j \cdot \mathrm{in}(c_j)$ intersects with f. If f is clear from the context, we omit it.

Let $I = (i_k)^4_{k=1}$ be $i_{lb}, i_{lu}, i_{rb}, i_{ru}$ ordered by size. Then I is called the *edge characteristics* of f. The *characteristics* of plane coloring f is a triple (a, b, I), where (a, b) is the frame of f, and I is the edge characteristics.

Proposition 5.4. *Let $(a, b, i_1, i_2, i_3, i_4)$ be the detailed frame of a plane coloring f. Then* $\mathrm{num}(f) \leqslant ab - \sum_{j=1}^4 \frac{i_j(i_j+1)}{2}$.

Definition 5.5 (*Diagonal caveat*). A *diagonal caveat* in f is a k-tuple of points $(\vec{p}_1, \ldots, \vec{p}_k)$ of D'_3 with $k \geqslant 3$ such that

$$\forall 1 \leqslant j < k\colon \quad \left(\vec{p_{j+1}} = \vec{p}_j + \begin{pmatrix}0\\1\\1\end{pmatrix}\right) \vee \forall 1 \leqslant j < k\colon \left(\vec{p_{j+1}} = \vec{p}_j + \begin{pmatrix}0\\1\\-1\end{pmatrix}\right),$$

$$f(\vec{p}_1) = 1 = f(\vec{p}_k),$$

$$\forall 1 < j < k\colon \quad f(\vec{p}_j) = 0.$$

The number of diagonal caveats in f is denoted by $\mathrm{diagcav}(f)$.

The next lemma gives us a good bound on the number of 3-points of a plane coloring f, given its edge characteristics. Recall the above example coloring f_{ex} with detailed frame $(6, 9, 3, 2, 1, 2)$. Since the coloring does not have any diagonal caveats, the next lemma will show that $\mathrm{xsteps}(f)$ is given by $3 + 2 + 1 + 2 = 8$, as we have indicated.

Lemma 5.6. *Let f be a connected, caveat-free coloring of the plane $x = c$ which has a detailed frame $(a, b, i_1, i_2, i_3, i_4)$. Then*

$$\mathrm{xsteps}(f) = \sum_{j\in[1..4]} i_j - \mathrm{diagcav}(f).$$

Proof. It is sufficient to prove the lemma for the special case that f has the detailed frame $(a, b, i_{lb}, 0, 0, 0)$. The reason is just that from any connected, caveat-free plane coloring f we can generated four colorings (f_1, f_2, f_3, f_4) with detailed frames $(a_1, b_1, i_{lb}, 0, 0, 0) \ldots (a_4, b_4, 0, 0, 0, i_{ru})$ such that

$$\#3(f) = \sum_{j\in[1..4]} \#3(f_j).$$

We prove the case $(a, b, i_{lb}, 0, 0, 0)$ by induction. The base cases $a = b = 1$, $a = 2, b = 1$ and $a = 2 = b$ are trivial. For the induction step, let f be a plane coloring with detailed frame $(a, b, i_{lb}, 0, 0, 0)$ such that $(a, b) \geqslant (2, 2)$. If $i_{lb} = 0$, then $\#3(f) = 0$ and $\mathrm{diagcav}(f) = 0$.

Otherwise let f' be generated from f by deleting the first column. I.e.,

$$f'(x, y, z) = \begin{cases} 0 & \text{if } y = \min_y(f), \\ f(x, y, z) & \text{else.} \end{cases}$$

Then

$$c^{f'}_{lb} = c^f_{lb} + \begin{pmatrix}0\\1\\0\end{pmatrix} = c^f_{lb} + \begin{pmatrix}0\\0.5\\-0.5\end{pmatrix} + \begin{pmatrix}0\\0.5\\0.5\end{pmatrix}.$$

Hence,

$$\begin{aligned}
\mathrm{ta}^{f'}\big(c^{f'}_{lb}\big) &= \mathrm{Lin}\begin{pmatrix}0\\0.5\\-0.5\end{pmatrix} + c^{f'}_{lb} \\
&= \mathrm{Lin}\begin{pmatrix}0\\0.5\\-0.5\end{pmatrix} + c^f_{lb} + \begin{pmatrix}0\\0.5\\-0.5\end{pmatrix} + \begin{pmatrix}0\\0.5\\0.5\end{pmatrix} \\
&= \mathrm{Lin}\begin{pmatrix}0\\0.5\\-0.5\end{pmatrix} + c^f_{lb} + \mathrm{in}^f\big(c^f_{lb}\big) \\
&= \mathrm{ta}^f\big(c^f_{lb}\big) + 1 \cdot \mathrm{in}^f\big(c^f_{lb}\big),
\end{aligned} \tag{9}$$

which implies that for any $k > 0$

$$\mathrm{ta}^{f'}\big(c^{f'}_{lb}\big) + (k-1) \cdot \mathrm{in}^{f'}\big(c^{f'}_{lb}\big) = \mathrm{ta}^f\big(c^f_{lb}\big) + k \cdot \mathrm{in}^f\big(c^f_{lb}\big). \tag{10}$$

Since for any k with $k < i_{lb}$, $\mathrm{ta}^f(c^f_{lb}) + k \cdot \mathrm{in}^f(c^f_{lb})$ does not intersect with f, we know that f' has a detailed frame $(a, b-1, i'_{lb}, 0, 0, 0)$ with $i'_{lb} \geqslant i_{lb} - 1$.

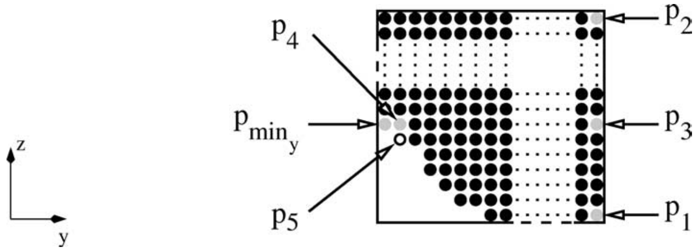


Fig. 9. Points considered in the proof of Lemma 5.6.

Let

$$\vec{p}_{min_y} = \begin{pmatrix} c \\ \min_y(f) \\ z \end{pmatrix}$$

be the point with

$$z = \min\left\{ z' \mid f \begin{pmatrix} c \\ \min_y(f) \\ z' \end{pmatrix} = 1 \right\}.$$

Since f has detailed frame $(a, b, i_{lb}, 0, 0, 0)$, we know that

$$\vec{p}_1 = \begin{pmatrix} c \\ \max_y(f) \\ \min_z(f) \end{pmatrix} \quad \text{and} \quad \vec{p}_2 = f \begin{pmatrix} c \\ \max_y(f) \\ \max_z(f) \end{pmatrix}$$

satisfy $f(\vec{p}_1) = 1 = f(\vec{p}_2)$. Since f is caveat-free, this implies that we have $f(\vec{p}_3) = 1$ for

$$\vec{p}_3 = \begin{pmatrix} c \\ \max_y(f) \\ z \end{pmatrix}.$$

Again since f is caveat-free, this implies that we have $f(\vec{p}_4) = 1$, where

$$\vec{p}_4 = \begin{pmatrix} c \\ \min_y(f) + 1 \\ z \end{pmatrix}.$$

Let

$$\vec{p}_5 = \begin{pmatrix} c \\ \min_y(f) + 1 \\ z - 1 \end{pmatrix} = \vec{p}_{min_y} + \begin{pmatrix} 0 \\ +1 \\ -1 \end{pmatrix}.$$

Fig. 9 shows the different points considered in the proof.

We distinguish the following two cases for the different colorings of point $\vec{p}_5$:

1. $f(\vec{p}_5) = 1$. Then

$$\text{diagcav}(f) = \text{diagcav}(f'). \tag{11}$$

Furthermore, either $\vec{p}_{min_y}$ is an element of $\mathrm{ta}^f(c^f_{lb}) + i_{lb} \cdot \mathrm{in}^f(c^f_{lb})$, which implies that $\vec{p}_5$ is an element of $\mathrm{ta}^f(c^f_{lb}) + i_{lb} \cdot \mathrm{in}^f(c^f_{lb})$, or $\vec{p}_{min_y}$ is not an element of $\mathrm{ta}^f(c^f_{lb}) + i_{lb} \cdot \mathrm{in}^f(c^f_{lb})$. In the latter case, there must be a $k > i_{lb}$ with $\vec{p}_{min_y} \in \mathrm{ta}^f(c^f_{lb}) + k \cdot \mathrm{in}^f(c^f_{lb})$. Then there must be another point $\vec{p}\,' = \begin{pmatrix} c \\ y \\ z \end{pmatrix}$ with $f(\vec{p}\,') = 1$, $y \neq \min_y(f)$ and $\vec{p}\,' \in \mathrm{ta}^f(c^f_{lb}) + i_{lb} \cdot \mathrm{in}^f(c^f_{lb})$. In both cases there is a point in $\mathrm{points}(f) \cap (\mathrm{ta}^f(c^f_{lb}) + i_{lb} \cdot \mathrm{in}^f(c^f_{lb}))$ with an y-coordinate different from $\min_y(f)$. Hence, this point is contained in f', which implies that this point is an element of $\mathrm{ta}^{f'}(c^{f'}_{lb}) + (i_{lb} - 1) \cdot \mathrm{in}^{f'}(c^{f'}_{lb})$ by Eq. (10). Then f' has detailed frame $(a, b-1, i_{lb}-1, 0, 0, 0)$. Since $(\vec{p}_5 - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \vec{p}_5, \vec{p}_{min_y})$ is an x-step in f but not in f', we get

$$\#3(f) = \#3(f') + 1.$$

Then

$$\begin{aligned} \#3(f) &= \#3(f') + 1 \\ &= \big[(i_{lb} - 1) - \mathrm{diagcav}(f')\big] + 1 \quad \text{(Ind. Hyp.)} \\ &= i_{lb} - \mathrm{diagcav}(f) \quad \text{(by (11))}. \end{aligned}$$

2. $f(\vec{p}_5) = 0$. Since there is no x-step between $\vec{p}_{min_y}$ and $\vec{p}_5$, we get

$$\#3(f) = \#3(f'). \tag{12}$$

We divide this case into two sub-cases:

(a) $\vec{p}_{min_y}$ is an element of a diagonal caveat of f: Then we know that there must be a point $\vec{p} \in \mathrm{ta}^f(c^f_{lb}) + i_{lb} \cdot \mathrm{in}^f(c^f_{lb})$ with $f(\vec{p}) = 1$ which has an y-coordinate different from $\min_y(f)$. Hence, f' has the detailed frame

$$(a, b-1, i_{lb} - 1, 0, 0, 0).$$

Since we have removed one diagonal caveat by deleting the first column (namely the one starting with $\vec{p}_{min_y}$), we get

$$\mathrm{diagcav}(f) - 1 = \mathrm{diagcav}(f'). \tag{13}$$

Then

$$\begin{aligned} \#3(f) &= \#3(f') \quad \text{by (12)} \\ &= (i_{lb} - 1) - \big(\mathrm{diagcav}(f')\big) \quad \text{(Ind. Hyp.)} \\ &= (i_{lb} - 1) - \big(\mathrm{diagcav}(f) - 1\big) \quad \text{by (13)} \\ &= i_{lb} - \mathrm{diagcav}(f). \end{aligned}$$

(b) $\vec{p}_{min_y}$ is not an element of a diagonal caveat of f: Then

$$\mathrm{diagcav}(f) = \mathrm{diagcav}(f'). \tag{14}$$

Furthermore, $\vec{p}_{min_y}$ must be the only element of $\mathrm{ta}^f(c^f_{lb}) + i_{lb} \cdot \mathrm{in}^f(c^f_{lb})$ colored by f. Then $\vec{p}_4$ is an element of $\mathrm{ta}^f(c^f_{lb}) + (i_{lb}+1) \cdot \mathrm{in}^f(c^f_{lb})$ and is colored black by f'. Hence, we know that f' has the detailed frame

$$(a, b-1, i_{lb}, 0, 0, 0).$$

Then

$$\begin{aligned}
\#3(f) &= \#3(f') \quad \text{by (12)}\\
&= i_{lb} - \mathrm{diagcav}(f') \quad \text{(Ind. Hyp.)}\\
&= i_{lb} - \mathrm{diagcav}(f) \quad \text{by (14).} \qquad \square
\end{aligned}$$

A first overall bound on $\mathrm{xsteps}(f)$ is given in the next proposition. This holds also for the pathological cases, which will be excluded later. A more precise bound will be given in the next section.

Proposition 5.7. *Let f be a caveat-free coloring of plane $x = c$ with frame (a, b). Then* $\mathrm{xsteps}(f) \leqslant 2(\min(a, b) - 1)$.

Proof. Let f be a coloring of plane $x = c$. We will first show that $\mathrm{xsteps}(f) \leqslant 2(a-1)$ by induction on a. For the base case let f be a coloring of height 1. Then $\mathrm{xsteps}(f) = 0$. For the induction step, let f be a plane coloring of height $a+1$. Let f' be f with the last row deleted. Then every x-step (p_1, p_2, p_3) in f' is also an x-step in f. On the other hand, an x-step (p_1, p_2, p_3) for f is an x-step for f' iff both p_2 and p_3 are not in the last row. Thus,

$$\begin{aligned}
&\mathrm{xsteps}(f)\\
&\quad = \mathrm{xsteps}(f')\\
&\qquad + \left| \left\{ (p_1, p_2, p_3) \;\middle|\; \begin{array}{l} (p_1, p_2, p_3) \text{ is an x-step for } f,\\ \exists y_2, y_3\colon\ p_2 = \begin{pmatrix} c \\ y_2 \\ \max_z(f) \end{pmatrix} \wedge p_3 = \begin{pmatrix} c \\ y_3 \\ \max_z(f) - 1 \end{pmatrix} \end{array} \right\} \right|\\
&\qquad + \left| \left\{ (p_1, p_2, p_3) \;\middle|\; \begin{array}{l} (p_1, p_2, p_3) \text{ is an x-step for } f,\\ \exists y_2, y_3\colon\ p_2 = \begin{pmatrix} c \\ y_2 \\ \max_z(f) - 1 \end{pmatrix} \wedge p_3 = \begin{pmatrix} c \\ y_3 \\ \max_z(f) \end{pmatrix} \end{array} \right\} \right|.
\end{aligned}$$

Let

$$\begin{aligned}
y^f_{m1} &= \min\bigl\{y \mid f\bigl(c, y, \max_z(f)\bigr) = 1\bigr\} - 1,\\
y^f_{m2} &= \max\bigl\{y \mid f\bigl(c, y, \max_z(f)\bigr) = 1\bigr\} + 1,\\
y^{f'}_{m1} &= \min\bigl\{y \mid f\bigl(c, y, \max_z(f) - 1\bigr) = 1\bigr\} - 1,\\
y^{f'}_{m2} &= \max\bigl\{y \mid f\bigl(c, y, \max_z(f) - 1\bigr) = 1\bigr\} + 1.
\end{aligned}$$

Then by the above said and the caveat-freeness of f, the only possibilities for p_1 in the x-steps that are in f but not in f' are

$$p_1^1 = \begin{pmatrix} c \\ y_{m1}^f \\ \max_z(f) \end{pmatrix}, \qquad p_1^2 = \begin{pmatrix} c \\ y_{m2}^f \\ \max_z(f) \end{pmatrix},$$
$$p_1^3 = \begin{pmatrix} c \\ y_{m1}^{f'} \\ \max_z(f) - 1 \end{pmatrix}, \qquad p_1^4 = \begin{pmatrix} c \\ y_{m2}^{f'} \\ \max_z(f) - 1 \end{pmatrix}.$$

We will show that if p_1^1 is contained in an x-step for f, then p_1^3 is not. The same holds for p_1^2 and p_1^4.

Now if there are points p_2^1 and p_3^1 such that (p_1^1, p_2^1, p_3^1) is an x-step in f but not in f', then

$$p_3^1 = \begin{pmatrix} c \\ y_{m1}^f \\ \max_z(f) - 1 \end{pmatrix},$$

which implies that $y_{m1}^{f'} < y_{m1}^f$. But then we get $f(c, y_{m1}^{f'}, \max_z(f)) = 0$ by the caveat-freeness of f, which implies that p_1^3 can not be part of an x-step that is in f but not in f'. Analogously, we get that if p_1^3 is part of an x-step that is in f but not in f', then p_1^1 is not. We get similar results for p_1^2 and p_1^4, which shows that we can add at most 2 x-steps in f. Thus, we have $\mathrm{xsteps}(f) \leqslant \mathrm{xsteps}(f') + 2$, which proves the claim by induction hypotheses.

Analogously, we get $\mathrm{xsteps}(f) \leqslant 2(b-1)$, which shows

$$\mathrm{xsteps}(f) \leqslant 2(\min(a, b) - 1). \qquad \square$$

6. Number of contacts

As already mentioned in Section 3, for every coloring f we need to distinguish between contacts, where both points are in the same layer, and contacts, where the two corresponding points are in successive layers. The first one are called *layer contacts of f* (denoted by LC_f^c), whereas the later ones are called *interlayer contacts*. Since we can split every coloring into a set of plane colorings, we define this notions for plane colorings.

6.1. Layer contacts

Let f be a coloring of plane $x = c$. Since all colored points of f are in plane $x = c$, we can define the *layer contacts* LC_f^c of f in the plane $x = c$ by $\mathrm{LC}_f^c = \mathrm{con}(f)$. We define $\mathrm{LC}_{n,a,b}$ to be the maximum of all LC_f^c with $\mathrm{num}(f) = n$, f has frame (a, b) and f is a coloring of some plane $x = c$.

Proposition 6.1. *Under assumption of caveat-free colorings,* $\mathrm{LC}_{n,a,b} = 2n - a - b$.

Proof. Let f be a coloring of an arbitrary plane $x = c$. If f is caveat-free, then the surface of f in the plane $x = c$ is $2a + 2b$. Now we know that each of the n points has 4 neighbors, which are either occupied by another point, or by a surface point. Hence, we get $4n = 2\mathrm{LC}_{n,a,b} + 2a + 2b$. $\square$

6.2. Interlayer contacts

Definition 6.2 (*Interlayer contacts*). Let f be a coloring of plane $x = c$, and f' be a coloring of plane $x = c'$. If $c' = c + 1$ (resp. $c - 1$), then we define the *interlayer contacts* $\mathrm{IC}_f^{f'}$ to be the number of contacts between plane $x = c$ and $x = c + 1$ (resp. $x = c - 1$) in the coloring $f \uplus f'$, i.e.:

$$\mathrm{IC}_f^{f'} = \left| \left\{ (\vec{p}, \vec{p}') \mid f(\vec{p}) = 1 \wedge f'(\vec{p}') = 1 \wedge \vec{p}' - \vec{p} = \begin{pmatrix} \pm 1 \\ \pm 0.5 \\ \pm 0.5 \end{pmatrix} \right\} \right|.$$

Otherwise, we define $\mathrm{IC}_f^{f'} = 0$.

Let f be a coloring of plane $x = c$. With $\mathrm{contacts}_{\max}(f, n)$ we denote the maximal number of contacts between plane $x = c$ and $x = c + 1$ by placing n points in $x = c + 1$. I.e.,

$$\mathrm{contacts}_{\max}(f, n) = \max \left\{ \mathrm{IC}_f^{f'} \,\middle|\, \begin{array}{l} f' \text{ is a plane coloring of } x = c + 1 \\ \text{with } \mathrm{num}(f') = n \end{array} \right\}.$$

Lemma 6.3. *Let f be a plane coloring of $x = c$. With $\delta_0(k)$ we denote* $\max(k, 0)$. *Then*

$$\begin{aligned} \mathrm{contacts}_{\max}(f, n) = 4 \min\bigl(n, \#4(f)\bigr) & \\ + 3 \min\bigl(\delta_0\bigl(n - \#4(f)\bigr), \#3(f)\bigr) & \\ + 2 \min\left(\delta_0\left(n - \sum_{i=3}^{4} \#i(f)\right), \#2(f)\right) & \\ + 1 \min\left(\delta_0\left(n - \sum_{i=2}^{4} \#i(f)\right), \#1(f)\right). & \end{aligned}$$

Proof. For the claim, it is sufficient to prove that every f' maximizing $\mathrm{IC}_f^{f'}$ satisfies if there is a k-point $\vec{p}$ with $k < 4$ and $f'(\vec{p}) = 1$, then all $k + 1$-points $\vec{p}'$ satisfy $f'(\vec{p}') = 1$. Now suppose that this would be not the case. Let f' be a coloring of plane $x = c + 1$ such that there is a k-point $\vec{p}$ with $f'(\vec{p}) = 1$, and that there is a $k + 1$-point $\vec{p}'$ with $f'(\vec{p}') = 0$.

Defining

$$f''(x,y,z) = \begin{cases} 1 & \text{if } \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \vec{p}, \\ 0 & \text{if } \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \vec{p'}, \\ f(x,y,z) & \text{else} \end{cases}$$

will give us an f'' with $\mathrm{num}(f'') = \mathrm{num}(f')$ and

$$\mathrm{IC}_f^{f''} = \mathrm{IC}_f^{f'} + 1,$$

which is a contradiction to our assumption. □

In addition, we want to show that it is sufficient to consider only plane colorings f which maximize #3(f). We will consider the case #r_not_overlaps(f) $= 0$ only. The case #r_not_overlaps(f) > 0 will be treated later.

Lemma 6.4. *Let f, f' be two plane colorings with frame (a,b),* $\mathrm{num}(f) = n = \mathrm{num}(f')$ *and* #r_not_overlaps(f) $= 0 =$ #r_not_overlaps(f') *such that* #3(f) $>$ #3(f'). *Then*

$$\forall n'\colon \quad \mathrm{contacts}_{\max}(f,n') \geqslant \mathrm{contacts}_{\max}(f',n').$$

Proof. Let f and f' be given as described. By Lemma 6.3, we know that the maximal interlayer contacts can be achieved by first occupying all 4-positions, then the 3-positions and so on. Let $l = \#3(f) - \#3(f')$. By Corollary 4.7, we know that

$$\#4(f) = \#4(f'),$$
$$\#3(f) = \#3(f') + l,$$
$$\#2(f) = \#2(f') - 2l,$$
$$\#1(f) = \#1(f') + l.$$

We consider the following cases for the number n' of colored points in the next layer:

1. $n' \leqslant \#4(f) + \#3(f) + \#2(f)$. Since we can color in f as many 4-points and 2-points as in f' but possibly more 3-points, we immediately get $\mathrm{contacts}_{\max}(f,n') \geqslant \mathrm{contacts}_{\max}(f',n')$.
2. $\#4(f) + \#3(f) + \#2(f) < n' \leqslant \#4(f) + \#3(f) + \#2(f) + l$. Let k be $n' - \#4(f) + \#3(f) + \#2(f)$. Then we have to color k 1-points for f, whereas we do not need to use 1-points for f' (where we can use 2-points instead). Thus, we loose k contacts here. Since $k \leqslant l$ and we gain l contacts by coloring l more 3-points in f than in f', we again get $\mathrm{contacts}_{\max}(f,n') \geqslant \mathrm{contacts}_{\max}(f',n')$.
3. $\#4(f) + \#3(f) + \#2(f) + l < n'$. In this case, we get

$$\mathrm{contacts}_{\max}(f,n') = \mathrm{contacts}_{\max}(f',n'). \qquad \square$$

Next, we want not to consider a special coloring, but only the frame the coloring has. With $\mathrm{MIC}^{n_2,a_2,b_2}_{n_1,a_1,b_1}$ we denote

$$\max \left| \left\{ \mathrm{IC}^{f_2}_{f_1} \,\middle|\, \begin{array}{l} \mathrm{num}(f_1) = n_1 \wedge \mathrm{frame}(f_1) = (a_1, b_1) \wedge \\ \mathrm{num}(f_2) = n_2 \wedge \mathrm{frame}(f_2) = (a_2, b_2) \end{array} \right\} \right|.$$

We define $\mathrm{MIC}^{n_2}_{n_1,a_1,b_1} = \max_{a_2,b_2} \mathrm{MIC}^{n_2,a_2,b_2}_{n_1,a_1,b_1}$.

Proposition 6.5.

$$\mathrm{MIC}^{n_2}_{n_1,a_1,b_1} = \max \left\{ \mathrm{contacts}_{\max}(f, n_2) \,\middle|\, \begin{array}{l} f \textit{ is a plane coloring} \\ \textit{with frame } (a_1, b_1) \\ \textit{and } \mathrm{num}(f) = n_1 \end{array} \right\}.$$

6.2.1. Normal colorings

Now we proceed as follows. We will first consider the case that the frame is sufficiently filled (where we define what this means in a moment). In this case, we can use $\mathrm{edge}(a, b, n)$ and $\mathrm{ext}(a, b, n)$ to bound the maximal number of x-steps (or 3-points) as described previously in Section 3. After that, we will show that we do not have to consider the frames which are not sufficiently filled (the pathological cases). We start with defining what “sufficiently filled” means.

Definition 6.6. Let a, b, n be positive numbers such that $ab \geqslant n$. We define $\mathrm{edge}(a, b, n)$ by

$$\mathrm{edge}(a, b, n) = \max\left\{k \in \mathbb{N} \mid ab - 4\frac{k(k+1)}{2} \geqslant n\right\}.$$

Let $k = \mathrm{edge}(a, b, n)$. Then we define

$$\mathrm{ext}(a, b, n) = \left\lfloor \frac{ab - 4\frac{k(k+1)}{2} - n}{k+1} \right\rfloor. \tag{15}$$

Intuitively, $\mathrm{edge}(a, b, n)$ is the lower bound for the indent from the corners of a coloring of n points with frame (a, b), if we try to make the indents as uniform as possible (since uniform indents generate the maximal number of x-steps). $\mathrm{ext}(a, b, n)$ is the number of times we can add 1 to $\mathrm{edge}(a, b, n)$. Note that (15) can be equivalently defined by

$$\mathrm{ext}(a, b, n) = \max\left\{r \in \mathbb{N} \mid ab - 4\frac{k(k+1)}{2} - r(k+1) \geqslant n\right\} \tag{16}$$

where $k = \mathrm{edge}(a, b, n)$.

Proposition 6.7. $0 \leqslant \mathrm{ext}(a, b, n) \leqslant 3$.

Proof. By contradiction. Let $k = \mathrm{edge}(a, b, n)$. Suppose that $\mathrm{ext}(a, b, n) \geqslant 4$. Then one would get

$$ab - 4\frac{k(k+1)}{2} - 4(k+1) \geqslant n,$$

$$ab - \frac{4k(k+1) + 8(k+1)}{2} \geqslant n,$$
$$ab - \frac{4(k+1)(k+2)}{2} \geqslant n.$$

But this would imply $\mathrm{edge}(a, b, n) \geqslant k + 1$, which is contradictory to our assumption that $k = \mathrm{edge}(a, b, n)$. □

Using this definitions, we can say what sufficiently filled means.

Definition 6.8 (*Normal*). Let n be an integer, (a, b) be a frame with $a \leqslant b$. Furthermore, let $k = \mathrm{edge}(a, b, n)$ and $r = \mathrm{ext}(a, b, n)$. We say that n is *normal for* (a, b) if either $4k + r < 2(a - 1)$, or $4k + r = 2(a - 1)$ and $ab - 4\frac{k(k+1)}{2} - r(k + 1) = n$.

The reason for using this notion is that if n is normal for (a, b), $\mathrm{edge}(a, b, n)$ and $\mathrm{ext}(a, b, n)$ yield a good bound on the number of x-steps of a plane coloring f. This will be shown in the next two lemmas.

Lemma 6.9. *If n is normal for (a, b) (with $a \leqslant b$), then there exists a caveat-free, connected plane coloring f such that* $\mathrm{xsteps}(f) = 4k + r$, *where* $k = \mathrm{edge}(a, b, n)$ *and* $r = \mathrm{ext}(a, b, n)$. *Furthermore, if $b \geqslant 3$, then this f satisfies* $\#\mathrm{r\_not\_overlaps}(f) = 0$.

The proof of this lemma is given in Appendix A.

Lemma 6.10. *Let (a, b) be a frame of a caveat-free and connected plane coloring f with $a \leqslant b$. Let* $k = \mathrm{edge}(a, b, \mathrm{num}(f))$ *and* $r = \mathrm{ext}(a, b, \mathrm{num}(f))$. *Then*

$$\mathrm{xsteps}(f) \leqslant \begin{cases} 4k + r & \textit{if } 4k + r < 2(a - 1), \\ 2(a - 1) & \textit{else}. \end{cases}$$

The proof of this lemma is given in Appendix A.

Definition 6.11 (*Upper bound for* $\mathrm{MIC}^{n'}_{n,a,b}$). Let n be a number and $a \leqslant b$ with $ab \geqslant n \geqslant \mathit{max}(a, b)$. Let $k = \mathrm{edge}(a, b, n)$ and $r = \mathrm{ext}(a, b, n)$, and let

$$l = \begin{cases} 4k + r & \text{if } 4k + r < 2(a - 1), \\ 2(a - 1) & \text{else}. \end{cases}$$

We define

$$\mathit{max}_4(a, b, n) = n + 1 - a - b, \qquad \mathit{max}_2(a, b, n) = 2a + 2b - 2l - 4,$$
$$\mathit{max}_3(a, b, n) = l, \qquad \mathit{max}_1(a, b, n) = l + 4.$$

With $\delta_0(n)$ we denote $\max(n, 0)$. Now we define

$$\begin{aligned} \mathrm{BMIC}^{n'}_{n,a,b} = {} & 4 \min\bigl(n', \mathit{max}_4(a, b, n)\bigr) \\ & + 3 \min\bigl(\delta_0\bigl(n' - \mathit{max}_4(a, b, n), \mathit{max}_3(a, b, n)\bigr)\bigr) \end{aligned}$$

$$+ 2\min\left(\delta_0\left(n' - \sum_{i=3}^{4} max_i(a,b,n)\right), max_2(a,b,n)\right)$$
$$+ 1\min\left(\delta_0\left(n' - \sum_{i=2}^{4} max_i(a,b,n)\right), max_1(a,b,n)\right).$$

Before we can prove that we can use $\mathrm{BMIC}_{n,a,b}^{n'}$ as an upper bound for $\mathrm{MIC}_{n,a,b}^{n'}$, we need to show that we can restrict ourself to plane colorings f with the property that $\#\mathrm{r\_not\_overlaps}(f) = 0$. To simplify matters (and since additionally we need them later), we introduce the concept of a line number distribution. A *line number distribution* is a function $D : \mathbb{Z} \to \mathbb{N}$ with the property that

$$\mathrm{dom}(D) = \{z \mid D(z) > 0\}$$

is finite. $\mathrm{dom}(D)$ is called the *domain* of D. The line number distribution D_f of a coloring f of the plane $x = c$ is defined by

$$D_f(z) = \left|\{y \mid f(c,y,z) = 1\}\right|.$$

Given D, we define $\mathrm{num}(D) = \sum_{i \in \mathrm{dom}(D)} D(i)$.

Lemma 6.12. *Let f be a connected coloring of plane $x = c$ with frame (a,b), $\mathrm{num}(f) = n$ and $\#\mathrm{r\_not\_overlaps}(f) > 0$. Then there is a f' with frame (a,b'), $\mathrm{num}(f') = n$ and $\#\mathrm{r\_not\_overlaps}(f) = 0$ such that*

$$b' \leqslant b,$$
$$D(f) = D(f'),$$
$$\forall n': \quad \mathrm{contacts}_{\max}(f',n') \geqslant \mathrm{contacts}_{\max}(f,n').$$

Proof. By induction. Let f be a coloring, and let z be a row such that we have $\mathrm{r\_overlap}^+(f,z) = 0$. Let f_1, f_2 with $f_1 \uplus f_2 = f$ be the sub colorings below (and including) z and above (including) $z+1$. Now we place f_1 above f_2 such that they have overlap of 1. Call this coloring f'. Then f' has height a and width b or $b-1$. Furthermore, we have

$$\#\mathrm{r\_not\_overlaps}(f') = \#\mathrm{r\_not\_overlaps}(f) - 1,$$
$$\mathrm{Surf}_{pl}(f') = \mathrm{Surf}_{pl}(f) - 2.$$

Thus, we have

$$\#4(f') = \#4(f).$$

Let D_f be the line number distribution associated to f. We have the following cases:

1. $D_f(z) = 1 = D_f(z+1)$. Then

$$\mathrm{xsteps}(f') = \mathrm{xsteps}(f) - 2.$$

By Lemma 4.6, we get

$$\#3(f') = \#3(f),$$
$$\#2(f') = \#2(f) + 1,$$
$$\#1(f') = \#1(f) - 2,$$

which gives us $\text{contacts}_{\max}(f', n') \geqslant \text{contacts}_{\max}(f', n')$ by Lemma 6.3.

2. $D_f(z) = 1 \wedge D_f(z+1) > 1$ or $D_f(z) > 1 \wedge D_f(z+1) = 1$. Then

$$\text{xsteps}(f') = \text{xsteps}(f) - 1.$$

By Lemma 4.6, we get

$$\#3(f') = \#3(f) + 1,$$
$$\#2(f') = \#2(f) - 1,$$
$$\#1(f') = \#1(f) - 1,$$

which gives us $\text{contacts}_{\max}(f', n') \geqslant \text{contacts}_{\max}(f', n')$ by Lemma 6.3.

3. $D_f(z) > 1 \wedge D_f(z+1) > 1$. Then

$$\text{xsteps}(f') = \text{xsteps}(f).$$

By Lemma 4.6, we get

$$\#3(f') = \#3(f) + 2,$$
$$\#\#2(f') = \#2(f) - 3,$$
$$\#\#1(f') = \#1(f),$$

which gives us $\text{contacts}_{\max}(f', n') \geqslant \text{contacts}_{\max}(f', n')$ by Lemma 6.3. □

Theorem 6.13. *Under the condition given in Definition* 6.11*, we get that* $\text{BMIC}^{n'}_{n,a,b}$ *is an upper bound for* $\text{MIC}^{n'}_{n,a,b}$*, i.e.,*

$$\forall a, b\ \exists b' \leqslant b: \quad \text{MIC}^{n'}_{n,a,b} \leqslant \text{BMIC}^{n'}_{n,a,b'}.$$

If n is normal for (a, b), then the above bound is tight, i.e., $\text{BMIC}^{n'}_{n,a,b} = \text{MIC}^{n'}_{n,a,b}$.

Proof. That there is a $b' \leqslant b$ such that $\text{BMIC}^{n'}_{n,a,b'}$ is an upper bound for $\text{MIC}^{n'}_{n,a,b}$ follows from Lemmas 4.6, 6.4, 6.12, 6.3, 6.10 and from the fact that all plane colorings f with frame (a, b) satisfy $\text{Surf}_{pl}(f) \geqslant 2a + 2b$. That the bound is tight if n is normal for (a, b) follows from Lemma 6.9. □

Note that any frame (a, b) for a connected, caveat-free coloring f with $\text{num}(f) = n$ will satisfy $ab \geqslant n \geqslant \max(a, b)$, which is the reason for the bound on n in the above definition. We need to investigate properties of frames with respect to normality in greater detail. The next lemma just states that normality is kept if we either add additional colored points without changing the frame, or we switch to a smaller frame for the same number of colored points.

Lemma 6.14. *Let n be normal for (a,b). Then all $n \leqslant n' \leqslant ab$ are normal for (a,b). Furthermore, for all (a',b') such that $a' \leqslant a \wedge b' \leqslant b$ with $a'b' \geqslant n$, we have n is normal for (a',b').*

The proof of this lemma is given in Appendix A.

Clearly, we want to search only through the normal frames in order to find the frame (a,b) which maximizes $\mathrm{MIC}_{n,a,b}^{n'}$, given n and n'. This will be subject of Theorem 6.16.

6.2.2. Restriction to normal colorings

For this purpose, we define colorings which have

- maximal number of x-steps for given frame (a,b) (i.e., $\mathrm{xsteps}(f) = 2(\min(a,b)-1)$),
- maximal number of colored points under the above restriction.

To achieve $\mathrm{xsteps}(f) = 2(\min(a,b)-1)$, we must have 2 x-steps in every line. By caveat-freeness, this implies that these maximal colorings are as given in Fig. 10.

The definition of these colorings is achieved by defining maximal line number distributions (where maximal refers to maximal x-steps). Line number distributions have been introduced earlier in Section 6.2.1. The important property of line number distributions is that one can easily obtain bounds on the maximal number of x-steps from the line number distribution of a coloring.

The maximal line number distribution for a frame (a,b) is given by $D_{\mathrm{max3}}^{a,b}$, which has the property that below the line with maximal number of colored points, we add 2 points from line to line, and after the maximal line we subtract 2 points. For every line number distribution D, we have defined a canonical coloring $f_{can(D)}$. $\mathrm{num}(D)$ is the number of colored points of D, which is the same as the points colored by $f_{can(D)}$. The precise definitions can be found in the appendix. Fig. 10 gives examples of the corresponding canonical colorings with maximal number of x-steps for the frames $(5,5)$, $(5,6)$, $(5,7)$ and $(6,7)$.

Now we want to find for a given n a minimal frame (a_m,b_m) such that (a_m,b_m) has maximal number of x-steps. For this purpose, we define a set of tuples

$$M = \bigcup \bigl\{ \{(n,n), (n,n+1), (n,n+2), (n+1,n+2)\} \mid n \text{ odd} \bigr\}.$$

Note that M is totally ordered by the lexicographic order on tuples. Hence, we can define $\mathrm{MinF}(n)$ to be the minimal element $(a,b) \in M$ such that $\mathrm{num}(D_{\mathrm{max3}}^{a,b}) \geqslant n$.

Note that we have excluded the case (n,n) with n even in the set M. The reason is that in this case, any coloring f of this frame which has maximal number of x-steps (namely

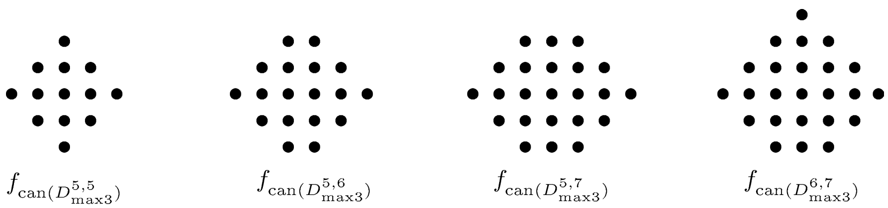


Fig. 10. Canonical colorings for the elements $(5,5)$, $(5,6)$, $(5,7)$ and $(6,7)$ of M.

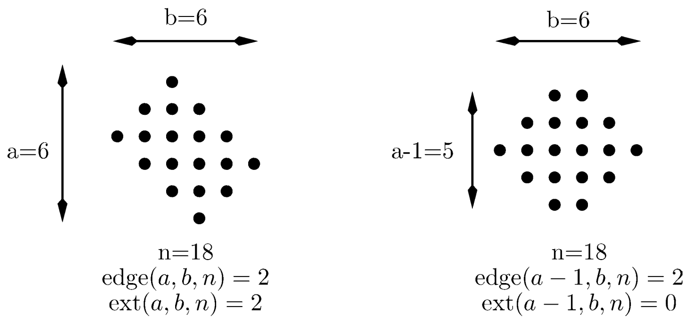


Fig. 11. Special case that a is even and $4\,\mathrm{edge}(a, b, n) + \mathrm{ext}(a, b, n) = 2(a-1)$. The first picture is the coloring for (a, b), the second for $(a-1, b)$.

$2(n-1))$ is not maximally overlapping. This implies that we can reduce this to a smaller frame. Fig. 11 shows an example.

Lemma 6.15. *Let n be a number and (a, b) be* $\mathrm{MinF}(n)$. *Then*

- *There is a plane coloring f with frame (a, b) such that* $\mathrm{num}(f) = n$ *and* $\mathrm{xsteps}(f) = 2(a-1)$.
- *n is normal for (a, b) or $(a, b-1)$.*

The proof is given in Appendix B.

Theorem 6.16 (Existence of optimal normal frame). *Let n be an integer. Then for all frames (a', b') there is a frame (a, b) such that $a \leqslant a' \wedge b \leqslant b'$, n is normal for (a, b), $(a, b-1)$ or $(a-1, b)$ and $\forall n'$:* $\mathrm{MIC}^{n'}_{n,a,b} \geqslant \mathrm{MIC}^{n'}_{n,a',b'}$.

Proof (*sketch*). The main idea of this theorem is the following. Fix n and n'. Let (a, b) be a frame for n with maximal number of possible x-steps (i.e., there is a plane coloring f with $\mathrm{num}(f) = n$, f has frame (a, b), and $\mathrm{xsteps}(f) = 2(\min(a, b) - 1)$). Then we know that $\mathrm{MIC}^{n'}_{n,a+1,b} \leqslant \mathrm{MIC}^{n'}_{n,a,b}$ since by enlarging the frame, we loose one 4-point by Lemma 4.6, but can win at most one x-step by Proposition 5.7. The same holds for $\mathrm{MIC}^{n'}_{n,a,b+1}$. Thus, it is sufficient to consider the minimal frame (a_m, b_m) which has maximal number of possible x-steps. But we can show that in this case, n is normal for (a_m, b_m), $(a_m, b_m - 1)$ or $(a_m - 1, b_m)$. $\square$

The full proof can be found in Appendix B.

This theorem states, that we need only to consider all frames that are within distance one from a normal frame in order to find the frame (a, b) with that maximizes $\mathrm{MIC}^{n'}_{n,a,b}$ for a given n and n'. Now we are able to summarize the results.

Theorem 6.17. *Let f be a connected, caveat-free coloring with $f = f_1 \uplus \cdots \uplus f_k$, where f_i is a coloring of the plane $x = i$. Then*

$$\mathrm{con}(f) \leqslant \sum_{i=1}^{k-1} \max \left\{ \mathrm{LC}_{n_i,a_i,b_i} + \mathrm{BMIC}^{n_{i+1}}_{n_i,a_i,b_i} \;\middle|\; \begin{array}{l} a_i b_i \geqslant n_i \text{ and } n_i \text{ is} \\ \text{normal for } (a_i, b_i), \\ (a_i - 1, b_i) \text{ or } (a_i, b_i - 1) \end{array} \right\} \tag{17}$$

$$+ \mathrm{LC}_{n_k,a_k^m,b_k^m}, \tag{18}$$

where $a_k^m = \lceil \sqrt{n_k} \rceil$ and $b_k^m = \lceil \frac{n_k}{a_k^m} \rceil$.

The proof is given in Appendix B.

7. Dynamic programming approach

Finally, we need an efficient method to calculate the bound given in Theorem 6.17. We apply an dynamic programming approach to calculate this bound. For this purpose, we define $B_1(n_1, n)$ to be an upper bound on the number of contacts for n colored points, provided that the first layer contains n_1 points. Formally, we define $B_1(n_1, n)$ recursively as follows:

$$\forall n: \quad B(n, n) = \mathrm{LC}_{n,a,b},$$

where $a = \lceil \sqrt{n} \rceil$ and $b = \lceil \frac{n}{a} \rceil$, and

$$\forall n \forall n_1 < n:$$
$$B(n_1, n) = \max_{\substack{1 \leqslant n_2 \leqslant n - n_1 \\ (a_1, b_1) \text{ frame for } n_1}} \left(\mathrm{LC}_{n_1,a_1,b_1} + \mathrm{BMIC}^{n_2}_{n_1,a_1,b_1} + B_1(n_2, n - n_1) \right),$$

where (a_1, b_1) is a frame for n_1 if $a_1 b_1 \geqslant n_1$ and n_1 is normal for (a_1, b_1), $(a_1 - 1, b_1)$ or $(a_1, b_1 - 1)$. Note that this implies that $a_1, b_1 \leqslant n_1$. Finally, we define

$$B(n) = \max_{1 \leqslant n_1 \leqslant n} B_1(n_1, n).$$

Proposition 7.1. *$B(n)$ can be calculated in $\mathrm{O}(n^2)$ space and $\mathrm{O}(n^5)$ time.*

Proposition 7.2. *For all $n_1 \leqslant n$, we have*

$$B_1(n_1, n) \geqslant \max \left\{ \mathrm{con}(f) \mid \exists k: \begin{pmatrix} f \text{ is connected and caveat-free} \\ \wedge f = f_1 \uplus \cdots \uplus f_k \\ \wedge \mathrm{num}(f) = n \wedge \mathrm{num}(f_1) = n_1 \end{pmatrix} \right\}.$$

Furthermore, $B(n)$ is an upper bound for the number of contacts $\mathrm{con}(f)$ *in any connected, caveat-free coloring f.*

Proof. Follows directly from Theorem 6.17. □

Table 1
Comparison of our bound with the previously introduced bound of $6n$ contacts

$n=?$	$B(n)$	$6n$
5	8	30
10	26	60
15	44	90
20	65	120
25	86	150
30	107	180
40	152	240
50	198	300
75	316	450
100	438	600
200	942	1200
300	1461	1800

Finally, we want to compare the bound yielded by our approach which the $6n$ bound that is used so far in the literature (e.g., in [1]). Table 1 shows a comparison of our bound with the $6n$ bound. The difference between our bound and the $6n$ bound is that our bound takes the surface of colorings into account, whereas the surface is ignored in the $6n$ bound. Since the surface grows slower with n than the number of contacts, it is clear that $B(n)$ asymptotically converges to $6n$.

8. Conclusion

We have presented an polynomial time upper bound for the number of contacts in the FCC-HP-model. The final upper bound is composed of an upper bound for the number of layer contacts, and an upper bound on the interlayer contacts.

There are two different outcomes of this research. The final bound $B(n)$ can be used in approximation algorithm (like [1]) to provide a sharper bound for the approximation ratio (at least for the case $n \leqslant 300$). The bounds on the layer and interlayer contacts on the other hand can be used in an branch-and-bound search for colorings that have maximal number of contacts for a given n. These colorings are called *hydrophobic cores*. They are important, since it seems to be easier to predict optimal conformations of an HP-chain by first predicting all optimal hydrophobic cores, and then try to thread the sequence on the hydrophobic cores. This could improve existing protein structure prediction approaches, where an FCC lattice model is used as an intermediate step [9,10].

Acknowledgement

I would like to thank Sebastian Will for many discussions on this topic, and for reading draft versions of this paper.

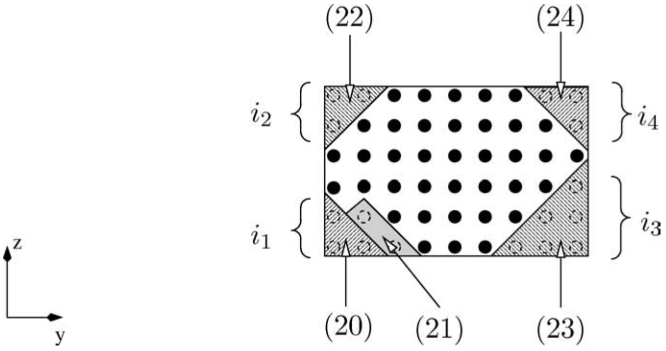


Fig. 12. Resulting coloring for Lemma 6.9. We have indicated the different regions defined in the proof. The numbers correspond to the formal definitions of these regions in the proof.

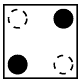
Appendix A. Proofs for Section 6.2.1

For Lemma 6.9, we have to show that if n is normal for a, b, then there is a coloring f such that xsteps(f) is 4edge(a, b, n) + ext(a, b, n). For this purpose, we start with a coloring f_{ab} that completely fills the frame (a, b). Then, we remove from f_{ab} diagonals from the edges such that the we have only n remaining points. Let $k = \text{edge}(a, b, n)$ and $r = \text{ext}(a, b, n)$. Define

$$\begin{aligned} i_1 &= k, & i_2 &= k + \delta_{r \geqslant 2}, \\ i_3 &= k + \delta_{r \geqslant 1}, & i_4 &= k + \delta_{r \geqslant 3}, \end{aligned} \tag{A.1}$$

where $\delta_{r \geqslant i}$ is 1 if $r \geqslant i$, and 0 otherwise. By this definition, we get $i_3 \geqslant i_2 \geqslant i_4 \geqslant i_1$. Then $i_1 \ldots i_4$ diagonals are removed from the corresponding edges (see Fig. 12), and the remaining $n_r = ab - 4\frac{k(k+1)}{2} - r(k+1) - n$ are removed from the bottom left corner. The tedious part in the following proof is to show that the excluded regions are actually disjoint, since otherwise we would exclude less points than required.

Furthermore, we have to show that the resulting coloring f satisfies #r_not_overlaps(f) $= 0$ if $b \geqslant 3$. Note that for the frame $(2, 2)$ and $n = 2$, this does not hold (albeit $n = 2$ is normal for $(2, 2)$). The resulting coloring f is of the form



and has #r_not_overlaps$(f) = 1$. For this case, we have two x-steps, but #3$(f) = 0$.

Proof of Lemma 6.9. Let n, a, b and k, r be given as defined in the lemma. Define f_{ab} by

$$f_{ab}(x, y, z) = \begin{cases} 1 & \text{if } x = 0,\ 1 \leqslant z \leqslant a,\ 1 \leqslant y \leqslant b, \\ 0 & \text{else.} \end{cases}$$

f_{ab} just fills the rectangle with side length a and b completely. The corners f_{ab} are

$$c_{lb} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad c_{lu} = \begin{pmatrix} 0 \\ 1 \\ a \end{pmatrix}, \quad c_{rb} = \begin{pmatrix} 0 \\ b \\ 1 \end{pmatrix}, \quad \text{and} \quad c_{ru} = \begin{pmatrix} 0 \\ a \\ b \end{pmatrix}.$$

Let $n_r = ab - 4\frac{k(k+1)}{2} - r(k+1) - n$. Then $0 \leqslant n_r < k+1$ by the definition of $r = \mathrm{ext}(a, b, n)$. Let $m = (a, b, i_1, i_2, i_3, i_4)$ be the tuple with $i_1 \dots i_4$ as defined by Eq. (A.1). We will show that there is a f with $m = (a, b, i_1, i_2, i_3, i_4)$ as a detailed frame.

Now define f by

$$\begin{aligned}
\mathrm{points}(f) = \mathrm{points}(f_{ab}) \\
&- \{\mathrm{ta}(c_{lb}) + l \cdot \mathrm{in}(c_{lb}) \mid 0 \leqslant l < i_1\} && \text{(A.2)}\\
&- \left\{ \begin{pmatrix} 0 \\ i_1+1 \\ 1 \end{pmatrix} + s \cdot \mathrm{tavec}(c_{lb}) \mid 0 \leqslant s < n_r \right\} && \text{(A.3)}\\
&- \{\mathrm{ta}(c_{lu}) + l \cdot \mathrm{in}(c_{lu}) \mid 0 \leqslant l < i_2\} && \text{(A.4)}\\
&- \{\mathrm{ta}(c_{rb}) + l \cdot \mathrm{in}(c_{rb}) \mid 0 \leqslant l < i_3\} && \text{(A.5)}\\
&- \{\mathrm{ta}(c_{ru}) + l \cdot \mathrm{in}(c_{ru}) \mid 0 \leqslant l < i_4\}. && \text{(A.6)}
\end{aligned}$$

See Fig. 12 for the location of the above defined regions. First, we have to show that the different exclusion sets are disjoint within the frame of f_{ab}, i.e., there is no point $\vec{p} = \begin{pmatrix} 0 \\ y \\ z \end{pmatrix}$ such that $1 \leqslant y \leqslant b$, $1 \leqslant z \leqslant a$ and $\vec{p}$ is in two of the exclusion sets.

For (A.2) and (A.3) it follows directly from the definition. Furthermore, we get that

$$\{\mathrm{ta}(c_{lb}) + l \cdot \mathrm{in}(c_{lb}) \mid 0 \leqslant l < i_1 + 1\} \tag{A.7}$$

contains both (A.2) and (A.3), and we will show that either (A.3) is empty, or that we will get pairwise disjointness of (A.7) with (A.4), (A.5) and (A.6).

So let's consider (A.5) and (A.6). Since $4k + r = i_3 + i_4 + i_2 + i_1 \leqslant 2(a-1)$, we get by definition of $m = (a, b, i_1, i_2, i_3, i_4)$ that

$$i_3 + i_4 \leqslant a - 1 \tag{A.8}$$

as follows: If $i_3 + i_4$ were greater than $a - 1$, then $i_3 + i_4 + i_2 + i_1 \geqslant a + a - 1 = 2a - 1 > 2(a-1)$ (since by definition of $m = (a, b, i_1, i_2, i_3, i_4)$ we know that $i_2 + i_1 \leqslant i_3 + i_4 \leqslant i_2 + i_1 + 1$), which would be a contradiction. Now let $\vec{p^3}_{maxz}$ be the point with $f_{ab}(\vec{p^3}_{maxz}) = 1$, $\vec{p^3}_{maxz}$ is contained in the set defined by (A.5), and has maximal z-value. By the definition of $\mathrm{ta}(c_{rb})$, $\vec{p^3}_{maxz}$ must have also maximal y-value. Now the maximal y-value that can be achieved in (A.5) is b. The z-value of a point in

$$\mathrm{ta}(c_{rb}) + l \cdot \mathrm{in}(c_{rb}) = \begin{pmatrix} 0 \\ b \\ 1 \end{pmatrix} + \mathrm{Lin}\begin{pmatrix} 0 \\ 0.5 \\ 0.5 \end{pmatrix} + l \cdot \begin{pmatrix} 0 \\ -0.5 \\ 0.5 \end{pmatrix}$$

which has y-value b is $1 + 2 \cdot 0.5 \cdot l = 1 + l$. Hence, we get

$$\vec{p^3}_{maxz} = \begin{pmatrix} 0 \\ b \\ 1 + (i_3 - 1) \end{pmatrix}.$$

Similarly, we define $\vec{p^4}_{minz}$ to be the point with $f_{ab}(\vec{p^4}_{minz}) = 1$, $\vec{p^4}_{minz}$ is contained in the set defined by (A.6), and has minimal z-value. Analogously, we get

$$\vec{p^4}_{minz} = \begin{pmatrix} 0 \\ b \\ a-(i_4-1) \end{pmatrix}$$

Now (A.5) and (A.6) would contain a common point if $1 + (i_3 - 1) = i_3 \geqslant a - (i_4 - 1)$, i.e., if $i_3 + i_4 \geqslant a + 1$, which is not the case by Eq. (A.8). By Eq. (A.8), we even get that the point $\begin{pmatrix} 0 \\ b \\ 1+i_3 \end{pmatrix}$ must be colored by f, which is a point in column $y = b$.

In the analogous prove for (A.2), (A.3) and (A.5), we get that two cases. Either $i_1 + i_2 + i_3 + i_4 = 2(a-1)$, in which case (A.3) is empty by the definition of "n normal for (a, b)", and we can adapt the above proof for (A.2) and (A.5). Or $i_1 + i_2 + i_3 + i_4 < 2(a-1)$, in which case we can conclude that $i_1 + i_3 < a - 1 \leqslant b - 1$ and we can adapt the above proof for (A.7) and (A.5) instead.

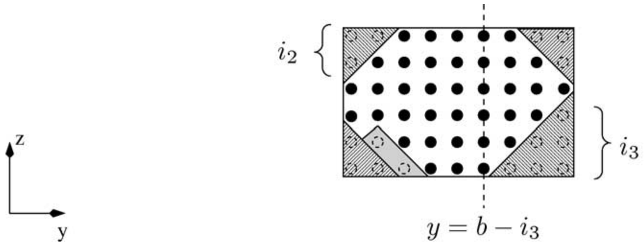
The cases (A.2) and (A.4), as well as (A.4) and (A.6) are analogous.

In any case, we will get that there are points colored by f in column $y = 1$ and $y = b$, and in the rows $z = 1$ and $z = a$. Hence, f has frame (a, b).

The remaining cases (A.2) and (A.6), as well as (A.4) and (A.5) are left to the reader.

Thus, we get that $\text{num}(f) = ab - 4\frac{k(k+1)}{2} - r(k+1) - n_r$, which is n. Furthermore, f has exactly $i_1 + i_2 + i_3 + i_4$ x-steps.

Finally, we have to show that $b \geqslant 3$ implies #r_not_overlaps$(f) = 0$. For this, we consider the column



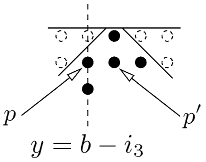
i.e., the set of points

$$C = \left\{ \begin{pmatrix} 0 \\ (b-i_3) \\ i \end{pmatrix} \,\middle|\, 1 \leqslant i \leqslant a \right\}.$$

We have two cases:

1. $f\begin{pmatrix} 0 \\ (b-i_3) \\ a \end{pmatrix} = 1$. By the definition of f, this implies that all points $\vec{p}$ in C satisfy $f(\vec{p}) = 1$, from which #r_not_overlaps$(f) = 0$ follows immediately.
2. $f\begin{pmatrix} 0 \\ (b-i_3) \\ a \end{pmatrix} = 0$. Since $i_4 \leqslant i_3$, we know that $\begin{pmatrix} 0 \\ (b-i_3) \\ a \end{pmatrix}$ can only be excluded by the left upper corner exclusion (A.4). This can be only the case if $i_2 + i_3 \geqslant b$. By Eqs. (A.1) and (A.8), we get $i_2 + i_3 \leqslant a \leqslant b$, which implies that the only possible case is $i_2 + i_3 = b$.

This implies by (A.1) that $i_2 = i_3$. Furthermore, we have already proven that there is a colored point in row $z = a$, which implies that $i_4 = i_2 - 1$. We have two subcases:
(a) $i_4 \geqslant 1$. Then we have the following situation:



By the definition of f, this implies that both $\vec{p} = \begin{pmatrix} 0 \\ (b-i_3) \\ a-1 \end{pmatrix}$ and $\vec{p}'$ are colored by f. This holds also for all points between $\vec{p}$ and $\begin{pmatrix} 0 \\ (b-i_3) \\ 1 \end{pmatrix}$, from which we conclude $\#\text{r\_not\_overlaps}(f) = 0$.

(b) $i_4 = 0$. In this case, i_3 and i_2 must be 1, which implies $b = 2$. $\square$

For Lemma 6.10, we have to show that for every plane coloring f with frame (a, b),

$$\text{xsteps}(f) \leqslant \begin{cases} 4k + r & \text{if } 4k + r < 2(a-1), \\ 2(a-1) & \text{else} \end{cases}$$

(where $k = \text{edge}(a, b, \text{num}(f))$ and $r = \text{ext}(a, b, \text{num}(f))$). For the case $4k + r < 2(a - 1)$, one has to show that the maximal number of x-steps can be achieved by distributing the edge indents $i_1 \dots i_4$ as uniformly as possible (i.e., such that $\forall j, j'$: $|i_j - i_{j'}| \leqslant 1$. This is done in the following proof.

Proof of Lemma 6.10. Let f, k and r be as given in the lemma. For the first case $4k + r < 2(a - 1)$, we define $\text{char}(i_1, i_2, i_3, i_4)$ to be the corresponding edge characteristics, i.e., the tuple generated by ordering i_1, i_2, i_3, i_4 by size. We define $\text{xsteps}(i_1, i_2, i_3, i_4)$ to be $i_1 + i_2 + i_3 + i_4$, and $\text{ex}(i_1, i_2, i_3, i_4)$ to be

$$\sum_{j=1}^{i_1} j + \sum_{j=1}^{i_2} j + \sum_{j=1}^{i_3} j + \sum_{j=1}^{i_4} j.$$

By Lemma 5.6, $\text{xsteps}(i_1, i_2, i_3, i_4)$ is the maximal number of x-steps that any caveat-free and connected plane coloring f with detailed frame $(a, b, i_1, i_2, i_3, i_4)$ can have. $\text{ex}(i_1, i_2, i_3, i_4)$ gives a bound on the number of points that may be colored by f by Proposition 5.4 (i.e., $ab - \text{ex}(i_1, i_2, i_3, i_4) \geqslant \text{num}(f)$). Furthermore, we define the ordering $\prec$ on quadruples by $(i_1, i_2, i_3, i_4) \prec (i'_1, i'_2, i'_3, i'_4)$ iff (i_1, i_2, i_3, i_4) is lexicographically smaller than (i'_1, i'_2, i'_3, i'_4).

Now define

$$I(i_1, i_2, i_3, i_4) = \big\{(i'_1, i'_2, i'_3, i'_4) \mid \text{xsteps}(i'_1, i'_2, i'_3, i'_4) = \text{xsteps}(i_1, i_2, i_3, i_4)\big\}.$$

Now if there is an element (i'_1, i'_2, i'_3, i'_4) of $I(i_1, i_2, i_3, i_4)$ such that there is a j and j' with $i'_j \leqslant i'_{j'} + 2$, then (i'_1, i'_2, i'_3, i'_4) is not maximal since substituting i'_j by $i'_j + 1$ and $i'_{j'}$ by

$i'_{j'} - 1$ will give an element $(i''_1, i''_2, i''_3, i''_4)$ such that

$$\begin{aligned}
&\mathrm{xsteps}(i'_1, i'_2, i'_3, i'_4) = \mathrm{xsteps}(i''_1, i''_2, i''_3, i''_4), \\
&\mathrm{char}(i'_1, i'_2, i'_3, i'_4) \prec \mathrm{char}(i''_1, i''_2, i''_3, i''_4), \\
&\mathrm{ex}(i'_1, i'_2, i'_3, i'_4) > \mathrm{ex}(i''_1, i''_2, i''_3, i''_4).
\end{aligned} \tag{A.9}$$

Hence, a $\prec$-maximal element (i'_1, i'_2, i'_3, i'_4) of $I(i_1, i_2, i_3, i_4)$ will satisfy $\forall j, j' : |i'_j - i'_{j'}| \leqslant 1$, and is minimal with respect to $\mathrm{ex}()$.

Now let f be a plane coloring of frame (a, b) such that $4k + r < 2(a - 1)$, and let f have the detailed frame $(a, b, i^f_1, i^f_2, i^f_3, i^f_4)$ with $i^f_1 + i^f_2 + i^f_3 + i^f_4$. Then

$$\mathrm{xsteps}(f) \leqslant i^f_1 + i^f_2 + i^f_3 + i^f_4$$

by Lemma 5.6.

Let (i_1, i_2, i_3, i_4) be a $\prec$-maximal element in $I(i^f_1, i^f_2, i^f_3, i^f_4)$. Let k_m be $\min(i^f_1, i^f_2, i^f_3, i^f_4)$. Then $\forall j$: $k_m \leqslant i_j \leqslant k_m + 1$ by the maximality of (i_1, i_2, i_3, i_4). Let $0 \leqslant r_m \leqslant 3$ be the number of times such that $i_j = k_m + 1$. Then

$$\begin{aligned}
\mathrm{num}(f) = n &\leqslant ab - \mathrm{ex}\bigl(i^f_1, i^f_2, i^f_3, i^f_4\bigr) \\
&\leqslant ab - \mathrm{ex}(i_1, i_2, i_3, i_4) \quad \text{by (A.9)} \\
&= ab - \left(\sum_{j=1}^{i_1} j + \sum_{j=1}^{i_2} j + \sum_{j=1}^{i_3} j + \sum_{j=1}^{i_4} j\right) \\
&= ab - 4\left(\sum_{j=1}^{k_m} j\right) - r_m(k_m + 1) \\
&= ab - 4\frac{k_m(k_m + 1)}{2} - r_m(k_m + 1).
\end{aligned}$$

Now this implies that $k_m \leqslant \mathrm{edge}(a, b, n)$ by definition of $\mathrm{edge}(a, b, n)$. If $k_m < \mathrm{edge}(a, b, n)$, then we get, $4k_m + r_m \leqslant 4\mathrm{edge}(a, b, n) + \mathrm{ext}(a, b, n)$ since $r_m \leqslant 3$. Otherwise, if $k_m = \mathrm{edge}(a, b, n)$, we get that $r_m \leqslant \mathrm{ext}(a, b, n)$ by the definition of $\mathrm{ext}(a, b, n)$. This implies

$$\mathrm{xsteps}(f) \leqslant \mathrm{xsteps}\bigl(i^f_1, i^f_2, i^f_3, i^f_4\bigr) = 4k_m + r_m \leqslant 4\mathrm{edge}(a, b, n) + \mathrm{ext}(a, b, n),$$

which proves the first case.

The second case follows from Proposition 5.7. □

Proof of Lemma 6.14. Without loss of generality, we can assume $a \leqslant b$. The first claim follows from the definition of normal. We will prove the second claim by induction.

We will prove the case for $a' = a - 1$ and $b' = b$. The proof for $a' = a$ and $b' = b - 1$ is analogous. For smaller frames, it follows by applying induction hypotheses.

Now let $a' = a - 1$ and $b' = b$. If $n < a'b'$, there is nothing to prove. Otherwise, if $n \geqslant a'b'$, let $k = \mathrm{edge}(a, b, n)$ and $r = \mathrm{ext}(a, b, n)$. Now the definition for k and r can be equivalently restated as follows. k and r are the uniquely determined integers such that

$0 \leqslant r \leqslant 3$

$$n + 4\sum_{j=1}^{k} j + (r+1)(k+1) > ab \geqslant n + 4\sum_{j=1}^{k} j + r(k+1). \tag{A.10}$$

Furthermore, let $k' = \mathrm{edge}(a', b', n)$ and $r' = \mathrm{ext}(a', b', n)$. We have to show the following:

Claim 1. $4k' + r' \leqslant 2(a'-1) = 2(a-1) - 2$.

Claim 2. *Either* $4k' + r' < 2(a-1) - 2$, *or* $4k' + r' = 2(a-1) - 2$ *and* $n + 4\sum_{j=1}^{k'} j + r'(k'+1) = a'b'$.

Since $4k + r \leqslant 2(a-1)$ by assumption, for proving Claim 1 it suffices to show that

$$4k' + r' \leqslant 4k + r - 2.$$

Given the above, then we know that

$$4k' + r = 2(a'-1) \Rightarrow 4k + r = 2(a-1). \tag{A.11}$$

Furthermore, we have

$$\begin{aligned} &4k + r = 2(a-1) \Rightarrow (r = 0 \ \vee r = 2) \\ \text{and} \quad &4k' + r' = 2(a'-1) \Rightarrow (r' = 0 \vee r' = 2). \end{aligned} \tag{A.12}$$

We have two cases:

1. $r \geqslant 1$. Now

$$\begin{aligned} &2(a-1) \geqslant 4k + 1, \\ &a - 1 \geqslant 2k + \frac{1}{2}, \\ &a \geqslant 2k + 1\frac{1}{2}, \\ &a \geqslant 2(k+1) \quad (a \text{ int.}) \end{aligned}$$

and henceforth $b \geqslant a \geqslant 2(k+1)$. By combining (A.10) for (a, b) and (a', b'), we get

$$\begin{gathered} n + 4\sum_{j=1}^{k} j + (r+1)(k+1) - b > (a-1)b \geqslant n + 4\sum_{j=1}^{k'} j + r'(k'+1) \\ \wedge\!\!\setminus \\ 4\sum_{j=1}^{k} j + (r+1)(k+1) - 2(k+1) > n + 4\sum_{j=1}^{k'} j + r'(k'+1). \end{gathered}$$

Since $r + 1 \leqslant 4$, we get immediately $k' \leqslant k$. We have two subcases:

(a) $k' = k$. Then

$$
\begin{aligned}
&n + 4\sum_{j=1}^{k} j + (r+1)(k+1) - 2(k+1) > n + 4\sum_{j=1}^{k} j + r'(k+1),\\
&(r+1)(k+1) - 2(k+1) > r'(k+1),\\
&r + 1 - 2 > r',
\end{aligned}
$$

which implies that $r \geqslant 2$ and $r' \leqslant r - 2$. Hence,

$$4k' + r' = 4k + r' \leqslant 4k + r - 2,$$

which shows the first claim.

For the second claim, if $4k' + r' < 2(a-) - 2$, then there is nothing to prove. Now assume that $4k' + r' = 2(a - 1) - 2$. Then we know by Eq. (A.11) that also $4k + r = 2(a - 1)$. Now $4k + r = 2(a - 1)$ implies $r = 2$ by Eq. (A.12) (since we have assumed $r \geqslant 1$). Furthermore, we know that $ab = n + 4\sum_{j=1}^{k} j + 2(k + 1)$ since n is normal for b. Since $k = k'$, $r = 2$ and $0 \leqslant r' \leqslant r - 2$, we get $r' = 0$. Hence,

$$a'b' = (a-1)b \geqslant n + 4\sum_{j=1}^{k} j$$

by Eq. (A.10) applied to a', b' and k', r'. Since by our assumptions $4k' + r' = 4k = 2(a - 1) - 2$, we have to show that $(a - 1)b \leqslant n + 4\sum_{j=1}^{k} j$. Now

$$
\begin{aligned}
(a-1)b &= n + 4\sum_{j=1}^{k} j + 2(k+1) - b\\
&\quad\ \rotatebox{90}{$\geqslant$}\\
(a-1)b &\leqslant n + 4\sum_{j=1}^{k} j + 2(k+1) - 2(k+1).
\end{aligned}
$$

Hence, $n + 4\sum_{j=1}^{k} j = n + 4\sum_{j=1}^{k'} j \geqslant (a - 1)b$, which proves Claim 2. Note that this implies that if $4k + r = 2(a - 1)$, then $4k' + r' = 2(a' - 1)$. Furthermore, we know that a is $2k + 2$, which implies that a is even. Fig. 11 shows an example of this special case.

(b) $k' \leqslant k - 1$. Then

$$
\begin{aligned}
4k' + r' &\leqslant 4(k-1) + r'\\
&\leqslant 4k - 4 + 3 = 4k - 1 \quad (r' \leqslant 3)\\
&\leqslant 4k + r - 2 \quad (r \geqslant 1).
\end{aligned}
$$

In this case, we have either $r' = 3$, which implies by (A.12) that $4k' + r' < 2(a' - 1)$, or $r' \leqslant 2$, which implies $4k' + r' < 4k + r - 2$. Again, this gives us $4k' + r' < 2(a' - 1)$, which proves the second claim for this case.

2. $r = 0$. Then

$$a \geqslant 2k + 1,$$

and therefore $b \geqslant a \geqslant 2k+1$. Now by Eq. (A.10) applied to a, b and a', b', we get

$$\begin{aligned}
&n + 4\sum_{j=1}^{k} j + 1(k+1) - b > (a-1)b \geqslant n + 4\sum_{j=1}^{k'} j + r'(k'+1) \\
&\qquad\wedge\!\!\backslash \\
&n + 4\sum_{j=1}^{k} j + 1(k+1) - 2k - 1 > n + 4\sum_{j=1}^{k'} j + r'(k'+1).
\end{aligned}$$

This gives immediately $k' \leqslant k$. Now if k' were the same as k, then we would get

$$n + 4\sum_{j=1}^{k} j - k > n + 4\sum_{j=1}^{k} j + r'(k+1)$$

which is a contradiction since $0 \leqslant r'$. Hence, we can conclude that $k' \leqslant k-1$. If $k' \leqslant k-2$, then $4k' + r' < 4k - 2$ follows immediately. Otherwise, if $k' = k-1$, then

$$\begin{aligned}
&n + 4\sum_{j=1}^{k} j - k > n + 4\sum_{j=1}^{k-1} j + r'k, \\
&4k - k > r'k, \\
&3 > r',
\end{aligned}$$

which implies $r \leqslant 2$. Therefore, $4k' + r' \leqslant 4(k-1) + 2 = 4k - 2$, which proves Claim 1. For Claim 2, if $4k' + r' < 2(a-1) - 2$, then there is nothing to prove. So assume that $4k' + r' = 2(a-1) - 2$. Then we know by Eq. (A.11) that also $4k + r = 2(a-1)$. Since n is normal for a, b, this implies that

$$ab = n + \sum_{j=1}^{k} j.$$

Now

$$\begin{aligned}
(a-1)b &\leqslant ab - 2k - 1 \quad (b \geqslant 2k+1) \\
&= n + \sum_{j=1}^{k} j - 2k - 1 \\
&= n + \sum_{j=1}^{k-1} j + 2k - 1.
\end{aligned}$$

This implies $r' < 2$ by Eq. (A.10) applied to $(a', b') = (a-1, b)$. This gives us $4k' + r' < 4k - 2$, and therefore a contradiction to our assumption that $4k' + r' = 2(a-1) - 2$. This proves Claim 2 for this case. □

Appendix B. Proofs for Section 6.2.2

We need some additional notions. Let D be a line number distribution. We say that D is *connected* if $\mathrm{dom}(D) = [\min \mathrm{dom}(D) \ldots \max \mathrm{dom}(D)]$. Let z_m^D be an element in $\mathrm{dom}(D)$ with $\forall z\colon\ D(z) \leqslant D(z_m^D)$. We say that D is *monotone* iff

$$\forall z \in \mathrm{dom}(D)\colon\ z \leqslant z_m^D \Rightarrow D(z-1) \leqslant D(z),$$
$$\forall z \in \mathrm{dom}(D)\colon\ z \geqslant z_m^D \Rightarrow D(z) \geqslant D(z+1).$$

We define the *canonical coloring* $f_{can(D)}$ inductively as follows. Let (y, z, n) be a triple of integers. Then

$$f_{y,z,n}(x, y', z') = \begin{cases} 1 & \text{if } x = 0,\ z = z' \text{ and } y \leqslant y' < y + n, \\ 0 & \text{const.} \end{cases}$$

$f_{y,z,n}$ is the coloring of row z, which starts at y and ends at $y + n - 1$ (i.e., has exactly n colored points). If $\mathrm{dom}(D) = \{d\}$, then

$$f_{can(D)} = f_{0,d,D(d)}.$$

Otherwise, let $m = \max \mathrm{dom}(D)$, and let D' be D except on m, where $D'(m) = 0$. Let $y_{D'}$ be the y-coordinate of the leftmost colored point of $f_{can(D')}$ in row $m - 1$. I.e.,

$$y_{D'} = \begin{cases} \min\{y \mid f_{can(D')}(0, y, m-1) = 1\} & \text{if } D(m-1) = D'(m-1) > 0, \\ 0 & \text{else.} \end{cases}$$

Then

$$f_{can(D)} = \begin{cases} f_{0,m,D(m)} \uplus f_{can(D')} & \text{if } D(m-1) = 0, \\ f_{y_{D'},m,D(m)} \uplus f_{can(D')} & \text{if } D(m-1) - 1 \leqslant D(m), \\ & \quad \text{and } D(m) \leqslant D(m-1) + 1, \\ f_{y_{D'}-1,m,D(m)} \uplus f_{can(D')} & \text{if } D(m) > D(m-1) + 1, \\ f_{y_{D'}+1,m,D(m)} \uplus f_{can(D')} & \text{else.} \end{cases}$$

Proposition B.1. *Let D be a line number distribution. Then*

$$\begin{aligned} &\mathrm{xsteps}(f_{can(D)}) \\ &\quad = 2\big|\big\{m \mid D(m) > 0 \wedge D(m+1) > 0 \wedge \big|D(m) - D(m+1)\big| \geqslant 2\big\}\big| \\ &\qquad + 1\big|\big\{m \mid D(m) > 0 \wedge D(m+1) > 0 \wedge \big|D(m) - D(m+1)\big| = 1\big\}\big|. \end{aligned}$$

Proposition B.2. *Let D be a line number distribution. If D is connected, then $f_{can(D)}$ is connected and satisfies*

$$\#\mathrm{r\_not\_overlaps}(f_{can(D)}) = 0.$$

If D is connected and monotone, then $f_{can(D)}$ is caveat-free.

Proposition B.3. *Let D be a connected, monotone line number distribution. Let $b = \max \mathrm{ran}(D)$ and $a = |\mathrm{dom}(D)|$. Then (a, b) is the frame of $f_{can(D)}$.*

In the following, we will consider line number distribution, whose canonical coloring has maximal number of x-steps within its frame, and cannot be extend without loosing an x-step.

Definition B.4 (*Maximal line number distribution*). Let (a, b) be a tuple with $a \leqslant b$ such that a is odd, or $a \neq b$. Then $D^{a,b}_{\max 3}$ is defined by

$$D^{a,b}_{\max 3}(z) = \begin{cases} b - 2|\lceil \frac{a}{2} \rceil - z| & \text{if } 1 \leqslant z \leqslant a, \\ 0 & \text{else.} \end{cases}$$

Note that we have excluded the case where we have a frame (a, a) with a even. The reason is just that in this case, any coloring f of this frame which has maximal number of x-steps (namely $2(a-1)$) is not maximally overlapping (i.e., there is a row z such that $\text{r\_overlap}^+(f, z) < |D_f(z) - D_f(z+1)|$). This implies that we can reduce this to a smaller frame. Fig. 11 shows an example.

Proposition B.5. *Let (a, b) and (a', b') be two frames with $a \leqslant b$ and $a' \leqslant b'$ such that*

$$a' < a \wedge b' \leqslant b \textit{ or } a' \leqslant a \wedge b' < b.$$

Then $\text{num}(D^{a',b'}_{\max 3}) < \text{num}(D^{a,b}_{\max 3})$.

Proposition B.6. $D^{a,b}_{\max 3}$ *is a line number distribution with* $\text{dom}(D) = [1..a]$. *Furthermore, it is monotone.*

Proof. By definition, we know that $D^{a,b}_{\max 3}$ is a function $D^{a,b}_{\max 3} : \mathbb{Z} \to \mathbb{N}$. We have to show that $D^{a,b}_{\max 3} \geqslant 1$ for every $1 \leqslant z \leqslant a$. Since $|\lceil \frac{a}{2} \rceil - z|$ is monotone decreasing in z from 1 to $\lceil \frac{a}{2} \rceil$, and monotone increasing in z from $\lceil \frac{a}{2} \rceil$ to a, it suffices to show that $D^{a,b}_{\max 3}(1) \geqslant 1$ and $D^{a,b}_{\max 3}(a) \geqslant 1$. For $D^{a,b}_{\max 3}(1)$, we have

$$\begin{aligned} D^{a,b}_{\max 3}(1) &= b - 2\left(\left\lceil \frac{a}{2} \right\rceil - 1\right) \geqslant b - 2\left(\frac{a+1}{2} - 1\right) \\ &= b - (a + 1 - 2) = b - (a - 1) \\ &\geqslant 1 \quad (a \leqslant b). \end{aligned}$$

For $D^{a,b}_{\max 3}(a)$, we have two cases:

1. a odd. Then

$$\begin{aligned} D^{a,b}_{\max 3}(a) &= b - 2\left(a - \left\lceil \frac{a}{2} \right\rceil\right) \\ &= b - 2a + 2\frac{a+1}{2} = b - a + 1 \\ &\geqslant 1 \quad (a \leqslant b). \end{aligned}$$

2. a even. Then we know that b is odd (since (a, a) with a even was excluded in the definition of the lemma). Hence, $b \geqslant a + 1$. Now

$$\begin{aligned} D^{a,b}_{\max 3}(a) &= b - 2\left(a - \left\lceil \frac{a}{2} \right\rceil\right) \\ &= b - 2a + 2\frac{a}{2} = b - a \geqslant 1 \quad (a + 1 \leqslant b). \qquad \square \end{aligned}$$

Now we want to find for a given n a minimal frame (a_m, b_m) such that (a_m, b_m) has maximal number of x-steps. For this purpose, we define a set of tuples

$$M = \bigcup \bigl\{\bigl\{(n, n), (n, n + 1), (n, n + 2), (n + 1, n + 2)\bigr\} \mid n \text{ odd}\bigr\}.$$

Note that M is totally ordered by the lexicographic order on tuples. Hence, we can define $\mathrm{MinF}(n)$ to be the minimal element $(a, b) \in M$ such that $\mathrm{num}(D^{a,b}_{\max 3}) \geqslant n$. Fig. 10 gives an example of the corresponding canonical colorings with maximal number of x-steps.

Proposition B.7. *Let (a, b) be a frame such that $D^{a,b}_{\max 3}$ is defined, and let n be a number. Then (a, b) is normal for n iff $\mathrm{num}(D^{a,b}_{\max 3}) \leqslant n$.*

Proof. If a is odd, then let k be $\lceil \frac{a}{2} \rceil - 1$. Then $2k + 1 = a$. Furthermore, we know that $D^{a,b}_{\max 3}(z) + 2 = D^{a,b}_{\max 3}(z + 1)$ for every line z with $1 \leqslant z \leqslant k = \lceil \frac{a}{2} \rceil - 1$. Similarly, we get $D^{a,b}_{\max 3}(z) - 2 = D^{a,b}_{\max 3}(z + 1)$ for every line z with $\lceil \frac{a}{2} \rceil = a - k \leqslant z < a$. Hence,

$$ab = \mathrm{num}\bigl(D^{a,b}_{\max 3}\bigr) + 2\sum_{j=1}^{k} j + 2\sum_{j=1}^{k} j.$$

This implies $k = \mathrm{edge}(b, \mathrm{num}(D^{a,b}_{\max 3}), a)$, $4k = 2(a - 1)$ and $ab = \mathrm{num}(D^{a,b}_{\max 3}) + 4\sum_{j=1}^{k} j$. This implies that $\mathrm{num}(D^{a,b}_{\max 3})$ is normal for (a, b). The rest follows from Lemma 6.14.

The case for a even is analogous. $\square$

Lemma B.8. *Let (a, b) with a odd and $a \leqslant b + 1$, or a even and $a \leqslant b + 2$. Then*

$$\forall 1 \leqslant z \leqslant a: \quad D^{a,b}_{\max 3}(z) = D^{a,b-1}_{\max 3}(z) - 1.$$

Furthermore, for any n such that $\mathrm{num}(D^{a,b-1}_{\max 3}) < n \leqslant \mathrm{num}(D^{a,b}_{\max 3})$, there is a connected, caveat-free coloring f such that $\mathrm{xsteps}(f) = 2(a - 1)$, $\mathrm{frame}(f) = (a, b)$ and $\mathrm{num}(f) = n$.

Proof. The first claim follows by the definition of $D^{a,b}_{\max 3}$ if $D^{a,b-1}_{\max 3}$ is defined, which is the case for all frames considered in the lemma.

For the second claim, we will construct a line number distribution D such that $f_{can(d)}$ has the required properties. By the first claim, we get $\mathrm{num}(D^{a,b}_{\max 3}(z)) - a = \mathrm{num}(D^{a,b-1}_{\max 3}(z))$, which implies by the definition of n that $\mathrm{num}(D^{a,b}_{\max 3}(z)) - n = d < a$.

- a odd. Let $d_1 = \lceil \frac{d}{2} \rceil$ and $d_2 = d - d_1$. Then $d_2 \leqslant d_1$. Since a is odd and $d < a$, we get
$$d_1 \leqslant \frac{a-1}{2} < \left\lceil \frac{a}{2} \right\rceil. \tag{B.1}$$
Furthermore,
$$\begin{aligned} a - d_2 + 1 &\geqslant a - d_1 + 1 \\ &\geqslant a - \frac{a-1}{2} + 1 = \frac{2a - a + 1 + 2}{2} > \left\lceil \frac{a}{2} \right\rceil. \end{aligned} \tag{B.2}$$
We define D by
$$D(z) = \begin{cases} D^{a,b}_{\max 3}(z) - 1 & \text{if } 1 \leqslant z \leqslant d_1, \\ D^{a,b}_{\max 3}(z) - 1 & \text{if } a - d_2 < z \leqslant a, \\ D^{a,b}_{\max 3}(z) & \text{else}. \end{cases}$$
By (B.1) and (B.2), D is well-defined and satisfies $\mathrm{num}(D) = n$. We have to show that $f_{can(D)}$ has frame (a, b). By the first claim, we get $\mathrm{dom}(D) = \mathrm{dom}(D^{a,b}_{\max 3}) = [1..a]$. Furthermore, (B.1) and (B.2) show $D(\lceil \frac{a}{2} \rceil) = D^{a,b}_{\max 3}(\lceil \frac{a}{2} \rceil) = b$. Since D is connected and monotone, this shows that $f_{can(D)}$ has frame (a, b) by Proposition B.3.
- a even. Let $d_1 = \lfloor \frac{d}{2} \rfloor$ and $d_2 = d - d_1$. Then $d_2 \geqslant d_1$. Since a even, we get
$$d_1 < \frac{a}{2} = \left\lceil \frac{a}{2} \right\rceil$$
and $d_2 \leqslant \frac{a}{2} = \lceil \frac{a}{2} \rceil$. Now,
$$a - d_2 + 1 \geqslant a - \frac{a}{2} + 1 > \left\lceil \frac{a}{2} \right\rceil.$$
Then we can proceed analogously to the previous case. □

Now we are able to proof Lemma 6.15.

Proof of Lemma 6.15.

1. a odd and $b = a$. Then the previous element in M is $(a - 1, b)$. We will show that $\mathrm{num}(D^{a,b}_{\max 3}) = \mathrm{num}(D^{a-1,b}_{\max 3}) + 1$, which implies $\mathrm{num}(D^{a,b}_{\max 3}) = n$ by the minimality of (a, b). Hence, $f_{can(D^{a,b}_{\max 3})}$ is the coloring we are looking for the first claim. Furthermore, n is normal for (a, b) by Proposition B.7, which shows the second claim.
To show $D^{a,b}_{\max 3} = D^{a-1,b}_{\max 3} + 1$, note that $\lceil \frac{a-1}{2} \rceil = \lceil \frac{a}{2} \rceil - 1$. Hence, we have for all $1 \leqslant z \leqslant a - 1$ that
$$\begin{aligned} D^{a-1,b}_{\max 3}(z) &= b - 2\left|\left\lceil \frac{a-1}{2} \right\rceil - z\right| = b - 2\left|\left\lceil \frac{a}{2} \right\rceil - 1 - z\right| \\ &= b - 2\left|\left\lceil \frac{a}{2} \right\rceil - (z+1)\right| = D^{a,b}_{\max 3}(z+1). \end{aligned}$$

Hence, $\mathrm{num}(D^{a,b}_{\max3}) - \mathrm{num}(D^{a-1,b}_{\max3}) = D^{a,b}_{\max3}(1)$. Now

$$\begin{aligned}D^{a,b}_{\max3}(1) &= a - 2\left(\left\lceil \frac{a}{2} \right\rceil - 1\right) \quad (a = b)\\ &= a - 2\left(\frac{a+1}{2} - 1\right) = a - (a+1) + 2 = 1 \quad (a \text{ odd}).\end{aligned}$$

2. a even and $b = a + 1$. Then the previous element in M is $(a-1, b)$. Since $\lceil \frac{a-1}{2} \rceil = \lceil \frac{a}{2} \rceil$, we get

$$\begin{aligned}D^{a-1,b}_{\max3}(z) &= b - 2\left|\left\lceil \frac{a-1}{2} \right\rceil - z\right|\\ &= b - 2\left|\left\lceil \frac{a}{2} \right\rceil - z\right| = D^{a,b}_{\max3}(z).\end{aligned}$$

Since $D^{a,b}_{\max3}(a) = b - 2(a - \lceil \frac{a}{2} \rceil) = b - a$ and $b = a + 1$, we get

$$\mathrm{num}\big(D^{a,b}_{\max3}\big) - \mathrm{num}\big(D^{a-1,b}_{\max3}\big) = D^{a,b}_{\max3}(a) = 1.$$

By the minimality of (a, b), this implies that $\mathrm{num}(D^{a,b}_{\max3}) = n$. Hence, $f_{can(D^{a,b}_{\max3})}$ is the coloring we are looking for the first claim. Furthermore, n is normal for (a, b) by Proposition B.7, which shows the second claim.

3. a odd, $b > a$. By the definition of $\mathrm{MinF}(n)$, we get that $\mathrm{num}(D^{a,b-1}_{\max3}) < n$. Hence, there exists a coloring f with $\mathrm{frame}(f) = (a, b)$, $\mathrm{num}(f) = n$ and $\mathrm{xsteps}(f) = 2(a - 1)$ by Lemma B.8, which shows the first claim. By Proposition B.7, we get that $\mathrm{num}(D^{a,b-1}_{\max3})$ is normal for $(a, b - 1)$. By Lemma 6.14, this implies that n is normal for $(a, b - 1)$, which proves the second claim. □

Finally, we can prove the main theorems.

Proof of Theorem 6.16. Without loss of generality, we can assume that $a' \leqslant b'$. Then we have to show that for any caveat-free, connected coloring f with frame (a', b'), there is a frame (a, b) and a coloring f_{ab} such that (a, b, f_{ab}) satisfies the following $Cond_{a',b',f}$:

$$\begin{aligned}&a \leqslant a' \wedge b \leqslant b'\\ &\wedge\ n \text{ normal for } (a, b), (a, b-1) \text{ or } (a-1, b)\\ &\wedge\big(\forall n'\colon\ \mathrm{contacts}_{\max}(f_{ab}, n') \geqslant \mathrm{contacts}_{\max}(f, n')\big).\end{aligned}$$

By Lemma 6.12, we can restrict ourself to colorings where $\#\mathrm{r\_not\_overlaps}(f) = 0$.

So let f be an arbitrary caveat-free, connected plane coloring that satisfies

$$\#\mathrm{r\_not\_overlaps}(f) = 0,$$

$\mathrm{num}(f) = n$ and $\mathrm{frame}(f) = (a', b')$. By Proposition 5.7, we know that $\mathrm{xsteps}(f) \leqslant 2(a' - 1)$. Furthermore, let (a_m, b_m) be $\mathrm{MinF}(n)$, and let f_m be the coloring as required by Lemma 6.15 for (a_m, b_m). Then (again by Lemma 6.15) $\mathrm{xsteps}(f_m) = 2(a_m - 1)$, and n is normal for (a_m, b_m) or $(a_m, b_m - 1)$. We have the following cases:

1. $a' > a_m$. We have the following subcases:
   (a) $b' \geqslant b_m$. Let $k = |a' - a_m|$ (note that $k > 0$). Since $a_m < a' \wedge b_m \leqslant b'$, we get $\mathrm{Surf}_{pl}(f_m) \leqslant \mathrm{Surf}_{pl}(f) - 2k$. Since $\mathrm{xsteps}(f) \leqslant 2(a' - 1)$ and $\mathrm{xsteps}(f_m) = 2(a_m - 1)$, we get by Lemma 4.6 that

$$\#4(f_m) \geqslant \#4(f) + k,$$
$$\#3(f_m) \geqslant \#3(f) - 2k,$$
$$\#2(f_m) \geqslant \#2(f) + 2k,$$
$$\#1(f_m) \geqslant \#1(f) - 2k.$$

   This implies $\forall n'$: $\mathrm{contacts}_{\max}(f_m, n') \geqslant \mathrm{contacts}_{\max}(f, n')$ by Lemma 6.3. By Lemma 6.15, n is normal for (a_m, b_m) or $(a_m, b_m - 1)$. Hence, (a_m, b_m, f_m) satisfies $Cond_{a',b',f}$.
   (b) $b' < b_m$. Since $b_m \leqslant a_m + 2$, we get

$$b' \leqslant b_m - 1 \leqslant a_m + 1 \leqslant a'.$$

   Since we have assumed $a' \leqslant b'$, the only case can be that $a' = b'$. Hence, $b_m - 1 = a_m + 1$, and therefore $b_m = a_m + 2$. By definition of M, this implies that a_m is odd, which implies that $a' = b' = a_m + 1$ is even.
   By Lemma 6.15, case 3 we get that n is normal for $(a_m, b_m - 1) = (a' - 1, b')$. Hence (a', b', f) satisfies condition $Cond_{a',b',f}$.
2. $a' = a_m$. We have the following subcases:
   (a) $b' \geqslant b_m$. Since $\mathrm{xsteps}(f) \leqslant 2(a_m - 1) = \mathrm{xsteps}(f_m)$ and $\mathrm{Surf}_{pl}(f_m) \leqslant \mathrm{Surf}_{pl}(f)$, we know that (a_m, b_m, f_m) satisfies the condition $Cond_{a',b',f}$.
   (b) $b' < b_m$. By Lemma 6.15, we get that n is normal for $(a_m, b_m) = (a', b_m)$ or $(a_m, b_m - 1) = (a', b_m - 1)$. Since $b' < b_m$, we get by Lemma 6.14 that n is also normal for (a', b'), which implies that (a', b', f) satisfies condition $Cond_{a',b',f}$.
3. $a' < a_m$. We have the following subcases:
   (a) $b' \leqslant b_m$. By Lemma 6.15, we get that n is normal for (a_m, b_m) or $(a_m, b_m - 1)$. Since $a' < a_m$ and $b' \leqslant b_m$, this implies that n is normal for (a', b') or $(a', b' - 1)$ by Lemma 6.14. Hence, (a', b', f) satisfies condition $Cond_{a',b',f}$.
   (b) $b' > b_m$. If n is normal for (a', b'), then (a', b', f) satisfies condition $Cond_{a',b',f}$. Otherwise, consider the frame (a', b_m). We have the following subcases:
      (i) $\mathrm{num}(D^{a',b_m}_{\max 3}) \leqslant n$. If $\mathrm{num}(D^{a',b_m}_{\max 3}) = n$, then $(a', b_m, f_{can(D^{a',b_m}_{\max 3})})$ satisfies condition $Cond_{a',b',f}$.
      So assume $D^{a',b_m}_{\max 3} < n$. Since $a' < a_m \leqslant b_m$, we know that $D^{a',b}_{\max 3}$ is defined for all $b \geqslant b_m$. Let $b_{max} \geqslant b_m$ be the maximal integer such that $\mathrm{num}(D^{a',b_{max}}_{\max 3}) \leqslant n$. By Proposition B.7, we know that n is normal for (a', b_{max}). By Lemma 6.14, we get that $b_{max} + 1 \leqslant b'$. Since $a' < a_m \leqslant b_m$ and $b_m < b_{max} + 1$, we can apply Lemma B.8 to n and $(a', b_{max} + 1)$. This lemma will give us a coloring f' such that hence, $(a', b_{max} + 1, f')$ satisfies the condition $Cond_{a',b',f}$.
      (ii) $\mathrm{num}(D^{a',b_m}_{\max 3}) > n$. Then $\mathrm{num}(D^{a_m,b_m}_{\max 3}) > \mathrm{num}(D^{a',b_m}_{\max 3})$ by Proposition B.5. We have shown in the proof of Lemma 6.15, cases 1 and 2, that $a_m = b_m$ or a_m

even implies $n = \mathrm{num}(D_{\max 3}^{a_m,b_m})$. Since we are in the sub-case $\mathrm{num}(D_{\max 3}^{a_m,b_m}) > n$, this implies by the definition of MinF() that $a_m \leqslant b_m - 1$. Since we have excluded the cases $a_m = b_m$ and a_m even, $a_m \leqslant b_m - 1$ implies that $(a_m, b_m - 1)$ is in M. This implies $\mathrm{num}(D_{\max 3}^{a_m,b_m-1}) < n$ by the minimality of (a_m, b_m). Since $a' < a_m$, we know that $D_{\max 3}^{a',b_m-1}$ is defined. By Proposition B.5, we get $\mathrm{num}(D_{\max 3}^{a',b_m-1}) < \mathrm{num}(D_{\max 3}^{a_m,b_m-1}) < n$. Since we have shown $a' < b_m - 1$ and $\mathrm{num}(D_{\max 3}^{a',b_m-1}) < n$, we can apply the proof of the previous case with $b_m - 1$ instead of b_m.

Proof of Theorem 6.17. By definition, we get

$$\mathrm{con}(f) = \sum_{i=1}^{k-1} \mathrm{IC}_{f_i}^{f_{i+1}} + \sum_{i=1}^{k} \mathrm{LC}_{f_i}^{c}.$$

Let (a_i^f, b_i^f) be the frame of f_i. Then we get by the definitions of $\mathrm{LC}_{n_i,a_i^f,b_i^f}$ and $\mathrm{MIC}_{n_i,a_i^f,b_i^f}^{n_{i+1}}$ that

$$\mathrm{MIC}_{n_i,a_i^f,b_i^f}^{n_{i+1}} \geqslant \mathrm{IC}_{f_i}^{f_{i+1}} \quad \text{and} \quad \mathrm{LC}_{n_i,a_i^f,b_i^f} \geqslant \mathrm{LC}_{f_i}^{c}.$$

This gives us

$$\begin{aligned}\mathrm{con}(f) \leqslant{} & \sum_{i=1}^{k-1} \max\bigl\{\mathrm{MIC}_{n_i,a_i,b_i}^{n_{i+1}} + \mathrm{LC}_{n_i,a_i,b_i} \mid a_i b_i \geqslant n_i\bigr\} \\ & + \max\{\mathrm{LC}_{n_k,a_k,b_k} \mid a_k b_k \geqslant n_k\}.\end{aligned}$$

By Proposition 6.1, we get

$$\forall (a_i', b_i') \forall (a_i, b_i): \quad a_i \leqslant a_i' \wedge b_i \leqslant b_i' \Rightarrow \mathrm{LC}_{n_i,a_i,b_i} \geqslant \mathrm{LC}_{n_i,a_i',b_i'}. \tag{B.3}$$

Thus, we get

$$\max\{\mathrm{LC}_{n_k,a_k,b_k} \mid a_k b_k \geqslant n_k\} = \mathrm{LC}_{n_k,a_k^m,b_k^m},$$

where $a_k^m = \lceil \sqrt{n_k} \rceil$ and $b_k^m = \lceil \frac{n_k}{a_k^m} \rceil$, which makes up the second summand (18) in the inequation of the theorem.

By the last theorem, we get

$$\begin{aligned}\forall (a_i', b_i') \exists (a_i, b_i): \quad & a_i \leqslant a_i' \wedge b_i \leqslant b_i' \\ & \wedge n_i \text{ is normal for } (a_i, b_i), (a_i, b_i - 1) \text{ or } (a_i - 1, b_i) \\ & \wedge \mathrm{MIC}_{n_i,a_i,b_i}^{n_{i+1}} \geqslant \mathrm{MIC}_{n_i,a_i',b_i'}^{n_{i+1}}.\end{aligned} \tag{B.4}$$

By Theorem 6.13 we get

$$\forall (a_i, b_i): \quad \mathrm{BMIC}_{n_i,a_i,b_i}^{n_{i+1}} \geqslant \mathrm{MIC}_{n_i,a_i,b_i}^{n_{i+1}}. \tag{B.5}$$

Eqs. (B.3), (B.4) and (B.5) together give us the first summand (17) of the inequation claimed in the theorem. □

References

[1] R. Agarwala, S. Batzoglou, V. Dancik, S.E. Decatur, M. Farach, S. Hannenhalli, S. Muthukrishnan, S. Skiena, Local rules for protein folding on a triangular lattice and generalized hydrophobicity in the HP-model, J. Comput. Biol. 4 (2) (1997) 275–296.
[2] R. Backofen, Constraint techniques for solving the protein structure prediction problem, in: M. Maher, J.-F. Puget (Eds.), Proceedings of 4th International Conference on Principle and Practice of Constraint Programming (CP'98), in: Lecture Notes in Computer Science, vol. 1520, Springer-Verlag, 1998, pp. 72–86.
[3] R. Backofen, Optimization techniques for the protein structure prediction problem, Habilitationsschrift, University of Munich, 1999.
[4] B. Berger, T. Leighton, Protein folding in the hydrophobic-hydrophilic (HP) modell is NP-complete, in: Proc. of the Second Annual International Conferences on Computational Molecular Biology (RECOMB98), New York, 1998, pp. 30–39.
[5] J.H. Conway, N.J.A. Sloane, Sphere Packings Lattices and Groups, third ed., Springer-Verlag, New York, 1996.
[6] P. Crescenzi, D. Goldman, C. Papadimitriou, A. Piccolboni, M. Yannakakis, On the complexity of protein folding, in: Proc. of STOC, 1998, pp. 597–603. Short version in: Proc. of RECOMB'98, pp. 61–62.
[7] W.E. Hart, S.C. Istrail, Fast protein folding in the hydrophobid-hydrophilic model within three-eighths of optimal, J. Comput. Biol. 3 (1) (1996) 53–96.
[8] K.F. Lau, K.A. Dill, A lattice statistical mechanics model of the conformational and sequence spaces of proteins, Macromolecules 22 (1989) 3986–3997.
[9] R. Samudrala, Y. Xia, M. Levitt, E.S. Huang, Ab initio prediction of protein structure using a combined hierarchical approach, Proteins: Structure, Function, and Genetics S3 (1999) 194–198.
[10] Y. Xia, E.S. Huang, M. Levittd, R. Samudrala, Ab initio construction of protein tertiary structures using a hierarchical approach, J. Mol. Biol. 300 (2000) 171–185.
[11] K. Yue, K.A. Dill, Forces of tertiary structural organization in globular proteins, Proc. Natl. Acad. Sci. USA 92 (1995) 146–150.